



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 110233

TO: Jeffrey Parkin
Location: cm1/8e15/8e12
Art Unit: 1648
Friday, December 12, 2003

Case Serial Number: 10/032162

From: Edward Hart
Location: Biotech-Chem Library
CM1-6B02
Phone: 305-9203

edward.hart@uspto.gov

Search Notes

Examiner Parkin,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart

THIS PAGE BLANK (US:70)

From: Parkin, Jeffrey
Sent: Wednesday, December 10, 2003 5:44 PM
To: STIC-Biotech/ChemLib
Subject: U.S. Serial No. 10/032,162

RECEIVED
DEC 11 2003

mej

Please search SEQ ID NOS.: 13, 15, and 17 from the aforementioned application (U.S. Serial No. 10/032,162) v. all relevant databases, including interference. Place results on BOTH disk and paper.
Thanks!

JSP
Au 1648
CM01-8E15
308-2227

13 AA 643
15 | 579
17 | 625

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 12/18/03
Date Completed: 12/18/03
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: 3
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: QSP
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (US 710)

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Query			Length	DB	ID	Description
	Score	Match	%				
1	3462	100.0	643	22	AAB61505	HIV-1 SOS gp140 gl	
2	3295	95.2	625	22	AAB61507	HIV-1 deltaV3 SOS	
3	3091.5	89.3	850	16	AAB87724	gp120 from the HIV	
4	3082	89.0	883	22	AAB82761	Ancestral HIV-1 gr	
5	3069.5	88.7	855	18	AAB11581	Human immunodefici	
6	3069.5	88.7	855	22	AAB89113	Env protein of the	
7	3056	88.3	879	22	AAB61506	HIV-1 deltaV1V2* S	
8	3041	87.8	847	21	AAV97073	Variant HIV-1 SR16	
9	3036	87.7	842	24	ABU66565	Human immunodefici	

ALIGNMENTS

REC'D
AAB61505

ID AAB61505 standard: Protein: 643 AA.

AA AAB61505:

XX
DT 05-APR-2001 (first entry)XX
DE
HTV-1 SOS am140 glycoprotein

XX anti-HTV: gene therapy

KW gp140. yy

OS Human immunodeficiency virus type 1.

PN WO200100648-A1.

PD 04-JAN-2001.

PF 23-JUN-2000; 2000WO-US17267.

PR 25-JUN-1999; 99US-0340992.

XX PA (PROG-) PROGENICS PHARM INC.

PA (AARO-) AARON DIAMOND AIDS RES CENT.
XX

PI Binley JM, Schuelke N, Olson WC, Maddon PJ, Moore JP;
XX

DR WPI; 2001-122993/13.
DR N-DCDB: AAF39591

XX
 DE
 Nov. 21, 1941
 1941

Db 343 WNNTEGNNTEGNTITLPCRIKQIINMWQEVGKAMYAPPPIRGQIRCSSNITGLLLTRDG 402
 Qy 421 GINENGTEIFRPGGDMRDNRSEFYKVKVKEIPLGVAPTKCKRRVVQREKRAVGIGAV 480
 Db 403 GINENGTEIFRPGGDMRDNRSEFYKVKVKEIPLGVAPTKCKRRVVQREKRAVGIGAV 462
 Qy 481 FLGFLGAAGSTMGAAASMTLTVOARLLLSGI VQOQNLLRAIEAQRMQLQTLVWGIKQLQA 540
 Db 463 FLGFLGAAGSTMGAAASMTLTVOARLLLSGI VQOQNLLRAIEAQRMQLQTLVWGIKQLQA 522
 Qy 541 RVLAVERYLGDQOLGIGCSGKLCCTAVPNASWSNKS LDRINWNTTWMEWEREIDNY 600
 Db 523 RVLAVERYLGDQOLGIGCSGKLCCTAVPNASWSNKS LDRINWNTTWMEWEREIDNY 582
 Qy 601 TSEIYTLIEESQNOQEKNEQELLELDKWASLWNNWFDITNWLWY 643
 Db 583 TSEIYTLIEESQNOQEKNEQELLELDKWASLWNNWFDITNWLWY 625
 ID AAR67724 standard; Protein; 850 AA.
 AC AAR67724;
 XX
 DT 25-MAR-2003 (updated)
 DT 07-SEP-1995 (first entry)
 XX
 DE gp120 from the HIV GNE8 isolate.
 XX
 KW HIV; human immunodeficiency virus; gp120; glycoprotein;
 KW GNE clone; GNE8 isolate.
 XX
 OS Human immunodeficiency virus type 1.
 XX
 PN W09428929-A1.
 XX
 PD 22-DEC-1994.
 XX
 PF 07-JUN-1994; 94WO-US06036.
 XX
 PR 07-JUN-1993; 93US-0072833.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Berman PW, Nakamura GR;
 XX
 WPI; 1995-036112/05.
 N-PSDB; AAQ76018.
 XX
 PT Use of HIV gp 120 polypeptide(s) - for developing probes for the
 PT analysis, prevention and therapy of HIV infection
 XX
 PS Claim 23; Page 27-30; 108pp; English.
 XX
 CC This protein is the gp120 (envelope glycoprotein) from the GNE8 isolate
 CC of HIV-1 (human immunodeficiency virus type 1). There are neutralising
 CC epitopes in the V2 and C4 domains of gp120, in addition to the
 CC neutralising epitopes in the V3 domain. Although the amino acid sequences
 CC of the neutralising epitopes in the V2, V4 and C4 domains are variable,
 CC the amount of variation is highly constrained. This facilitates the
 CC design of HIV subunit vaccines that can induce antibodies that neutralise
 CC the most common HIV strains for a given geographic region. This invention
 CC provides a multivalent gp120 subunit vaccine where the gp120 present in
 CC the vaccine is from at least two HIV isolates which have different amino
 CC acid sequences for a neutralising epitope in these regions.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 850 AA;
 Query Match 89.3%; Score 3091.5; DB 16; Length 850;
 Best Local Similarity 89.7%; Pred. No. 2.6e-152;

Matches 581; Conservative 26; Mismatches 32; Indels 9; Gaps 5;
 Qy 2 EKLWTVYGVVWKEATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENVTE 61
 Db 31 EKLWTVYGVVWKEATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEIGLENVTE 90
 Qy 62 HFNWKNMNVQMOQEDIIISLWDQSLKPCVKLTPLCVTLNCKDV-NATNTTNDSEGMERG 120
 Db 91 NFNWKNMNVQMOQEDIIISLWDQSLKPCVKLTPLCVTLNCKDV-NATNTTNDSEGMERG 150
 Qy 121 EIKKCSNITTSIRDEVOKEYALFKLDVVDIDNNNTSYRLISCDTSVITQACPKISEP 180
 Db 151 EIKKCSNITTSIRDKMNEYALFKLDVVDIDNDNTSYRLISCDTSVITQACPKVSEP 210
 Qy 181 IPIHYCAPAGPAILKCKDKFNGKPGQVQVSTVCTHGIRPVVSTQLLNLSLAEEVVI 240
 Db 211 IPIHYCAPAGPAILKCKDKFNGKPGQVQVSTVCTHGIRPVVSTQLLNLSLAEEVVI 270
 Qy 241 RSDNFTNNAKTIIVOLKESVEINCTRNPNNTKSIHIGPGRAFYTGTGIIIGDIRQAHCN 300
 Db 271 RSANFSDNAKTIIVOLKESVEINCTRNPNNTKSIHIGPGRAFYTGTGIIIGDIRQAHCN 330
 Qy 301 SRAKNWDTLKOIVIKLREOFENKTIIVFNHSSGGDPEIVMHSFNCEGEFFYCNSTLPNST 360
 Db 331 SSTKNWNTLKOIVTKLREHF-NKTIIVFNHSSGGDPEIVMHSFNCEGEFFYCNSTLPNST 389
 Qy 361 WN-----NNTGSSNTEGNTITLPCRIKQIINMWQEVGKAMYAPPPIRGQIRCSSNITGL 415
 Db 390 WNYTYTWNTEGSDNT-GRNITLQCRINQIINMWQEVGKAMYAPPPIRGQIRCSSNITGL 448
 Qy 416 LTRDGINENGTEIFRPGGDMRDNRSEFYKVKVKEIPLGVAPTKCKRRVVQREKRAV 475
 Db 449 LTRDGG-NNSETEIFRPGGDMRDNRSEFYKVKVKEIPLGVAPTKCKRRVVQREKRAV 507
 Qy 476 GIGAVFLGFLGAAGSTMGAAASMTLTVOARLLLSGI VQOQNLLRAIEAQRMQLQTLVWGI 535
 Db 508 GIGAVFLGFLGAAGSTMGAAASMTLTVOARLLLSGI VQOQNLLRAIEAQRMQLQTLVWGI 567
 Qy 536 KOLQARVLAVERYLGDQOLGIGCSGKLCCTAVPNASWSNKS LDRINWNTTWMEWER 595
 Db 568 KOLQARVLAVERYLGDQOLGIGCSGKLCCTAVPNASWSNKS LDRINWNTTWMEWER 627
 Qy 596 EIDNYTSEIYTLIEESQNOQEKNEQELLELDKWASLWNNWFDITNWLWY 643
 Db 628 EIDNYTSLIYSLIEESQNOQEKNEQELLELDKWASLWNNWFDITNWLWY 675
 RESULT 4
 AAB82761
 ID AAB82761 standard; Protein; 883 AA.
 AC AAB82761;
 XX
 DT 29-OCT-2001 (first entry)
 XX
 DE Ancestral HIV-1 group M, subtype B gp160 protein.
 KW HIV-1; env gene; gp160 protein; ancestral gene; phylogeny; vaccine.
 OS Human immunodeficiency virus type 1.
 PN WO200160838-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 16-FEB-2001; 2001WO-US05288.
 XX
 PR 18-FEB-2000; 2000US-0183659.
 XX
 PA (UNIW) UNIV WASHINGTON.
 XX
 PI Mullins JI, Rodrigo AG, Learn GH, Li F;
 XX

DR MPI; 2001-536565/59.
 XX N-PSDB; AAH26468.
 PT Preparing an ancestral viral amino acid sequence useful as a vaccine
 PT comprises determining a recent ancestor of a circulating virus by
 PT maximum likelihood phylogeny analysis -
 XX
 PS Claim 8; Page 54; 89pp; English.
 XX
 CC The present sequence is that of an ancestral HIV-1 group M, subtype
 CC B gp160 (env gene product) sequence. The invention provides
 CC compositions and methods for determining ancestral viral gene
 CC sequences and ancestral viral protein sequences for highly diverse
 CC viruses, such as HIV-1. The methods use samples of circulating
 CC viruses to determine an ancestral viral sequence by maximum
 CC likelihood phylogeny analysis. In the present case, the ancestral
 CC HIV-1 subtype B env sequence (see AAH26468) was determined using 38
 CC subtype B sequences (obtained from 9 different countries) and 3
 CC subtype D (outgroup) sequences. The distances between this ancestral
 CC viral sequence and circulating strains used to determine it were on
 CC average 12.3% (range: 8.0-21.0%) while the available specimens were
 CC 17.3% different from each other (range: 13.3-23.2%). Thus, the
 CC ancestor sequence was, on average, more closely related to any given
 CC circulating virus than to any other variant. The ancestral gp160
 CC sequence included a wide variety of immunogenically active peptides
 CC when processed for antigen presentation; nearly all known subtype B
 CC CTL epitope consensus amino acids were represented. Thus, an
 CC immunogenic composition to this subtype B ancestor protein will
 CC elicit broad neutralising antibody against HIV-1 isolates of the
 CC same subtype, and will also elicit a broad cellular response
 CC mediated by antigen-specific T-cells. A claimed vaccine composition
 CC comprises a viral ancestor protein or its immunogenic fragment,
 CC especially one derived from the HIV-1 group M subtype B gp160
 CC ancestral protein.
 XX
 SQ Sequence 883 AA;

Query Match 89.0%; Score 3082; DB 22; Length 883;
 Best Local Similarity 87.2%; Pred. No. 8.2e-152;
 Matches 591; Conservative 19; Mismatches 32; Indels 36; Gaps 9;

QY 2 EKLWTVYVYGVVWKEATTLFCASDAKAYDTEVHNWATHACVPTDPNPQEVLENVTE 61
 DB 1 EKLWTVYVYGVVWKEATTLFCASDAKAYDTEVHNWATHACVPTDPNPQEVLENVTE 90
 QY 62 HFNWKNWYEQHEDIISLWDSKPCVKLTPLCVTLNCKD---VNATNTDSE---- 114
 DB 1 HFNWKNWYEQHEDIISLWDSKPCVKLTPLCVTLNCKD---VNATNTDSE---- 114
 QY 91 NFNWKNWYEQHEDIISLWDSKPCVKLTPLCVTLNCKD---VNATNTDSE---- 150
 DB 1 NFNWKNWYEQHEDIISLWDSKPCVKLTPLCVTLNCKD---VNATNTDSE---- 150
 QY 115 -----GTM--ERGEIKNCSENIITTSIRDEVOKEVALFYKLDVVPID-----NNNTSYRL 161
 DB 1 TSSGGGFMGEKEGEIKNCSENIITTSIRDEVOKEVALFYKLDVVPID-----NNNTSYRL 210
 QY 162 ISCDTSVITQACPKISPEPIPIHYCAPAGFAILKCNCKTFNGKGPCKNVSTVQCTHGIRP 221
 DB 1 INCNTSVITQACPKVSEPIPIHYCAPAGFAILKCNCKTFNGKGPCKNVSTVQCTHGIRP 270
 QY 222 VVSTOLLNGLAEVEVIRSDNFTNNAKTIIVOLKESVEINCTRNPNNTKSIHIGPGR 281
 DB 1 VVSTOLLNGLAEVEVIRSDNFTNNAKTIIVOLKESVEINCTRNPNNTKSIHIGPGR 330
 QY 282 AFTYTGIIIGDIQAHNCISRAKNWDTLKOIVKLREQFENK--TIVFNHSSGGDPEIVM 339
 DB 1 ALVATGIIIGDIQAHNCISRAKNWDTLKOIVKLREQFENK--TIVFNHSSGGDPEIVM 390
 QY 340 HSFNCEGEFFYCNSTOLFNSWT-----NNNTEGSNNT--EGNTITLPCRIKQIINMWQ 390
 DB 1 HSFNCEGEFFYCNSTOLFNSWT-----NNNTEGSNNT--EGNTITLPCRIKQIINMWQ 450
 QY 391 EVGKAMYAPPIRGQIRCSNITGLLLTRDGGINEN-----GTEIFRPGGGMDRNRSEFY 446
 DB 1 EVGKAMYAPPIRGQIRCSNITGLLLTRDGGINEN-----GTEIFRPGGGMDRNRSEFY 510

QY 447 KKKVVKIEPLGVAPTCKKRRVVQREKRVG-ICAVFLGFLGAGSTMGASMTLTVOARL 505
 DB 1 KKKVVKIEPLGVAPTCKKRRVVQREKRVG-ICAVFLGFLGAGSTMGASMTLTVOARL 570
 QY 506 LLSGIVQOQNLLRAIEAQQRMQLTVMGIGIKQARVLAVERYLGDQQLLGIWCGSGKLI 565
 DB 1 LLSGIVQOQNLLRAIEAQQRMQLTVMGIGIKQARVLAVERYLGDQQLLGIWCGSGKLI 630
 QY 566 CTAVPWNASWNSKSLDRINNNMTWMEWEREDNTYSETYLTIEESQOQKNEQELLEL 625
 DB 1 CTAVPWNASWNSKSLDRINNNMTWMEWEREDNTYSETYLTIEESQOQKNEQELLEL 690
 QY 626 DKWASLWNNWFDITNWLWY 643
 DB 1 DKWASLWNNWFDITNWLWY 708
 QY 691 DKWASLWNNWFDITNWLWY 708
 DB 1 DKWASLWNNWFDITNWLWY 708

RESULT 5
 AAW11581
 ID AAW11581 standard; Protein; 855 AA.
 XX AAW11581;
 AC AAW11581;
 XX
 XX 25-MAR-2003 (updated)
 DT 25-MAR-1997 (first entry)
 DT
 DE Human Immunodeficiency Virus-1 strain BA-L envelope protein.
 XX
 XX Acquired immune deficiency syndrome; AIDS; envelope protein;
 KW env gene; vaccine.
 KW
 XX Human immunodeficiency virus type 1 (strain BA-L).
 OS
 XX US5576000-A.
 FN
 XX 19-NOV-1996.
 PD
 XX 15-FEB-1995; 95US-0388809.
 PF
 XX 17-OCT-1990; 90US-0599491.
 PR 25-FEB-1993; 93US-0022835.
 PR 15-FEB-1995; 95US-0388809.
 XX
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA
 XX Franchini G, Gallo RC, Gartner S, Lori FC, Markham PD;
 PI Popovic M, Reitz MS;
 XX WPI; 1997-011206/01.
 DR N-PSDB; AAT58551.
 DR
 XX New isolated envelope protein of HIV-1 strain BA-L and recombinant
 PT equivalents - useful as immunogens for vaccines and antibody prodn.,
 PT typical of US clinical isolates
 PT
 XX Claim 1; Fig 9; 86pp; English.
 PS
 XX A HindIII fragment of unintegrated viral DNA representing the BA-L
 CC genome was cloned into lambda phage Charon 28 DNA from total DNA of
 CC peripheral blood lymphocytes infected with and producing HIV-1 (BA-L).
 CC A positive clone was selected by hybridisation using a HIV-1 env
 CC probe. This clone, designated BA-Li, was found to contain the
 CC entire gene for the envelope protein on a 2.8 kb HindIII-XbaI
 CC fragment and a 0.4 kb EcoRI-HindIII fragment. When cloned together
 CC these fragments comprise the env gene, as well as the coding regions
 CC for rev and the rev-responsive element of env, both necessary for
 CC efficient expression in eukaryotic cells. The claimed recombinantly
 CC produced envelope protein can be used as an immunogen for raising
 CC antibodies against HIV
 CC (Updated on 25-MAR-2003 to correct PF field.)
 CC
 XX Sequence 855 AA;

Query Match 88.7%; Score 3069.5; DB 18; Length 855;
 Best Local Similarity 88.1%; Pred. No. 3.5e-151;
 Matches 572; Conservative 27; Mismatches 43; Indels 7; Gaps 2;

QY 2 EKLWTVYGVVPVWKEATTLFCASDAKAYDETEVHNWATHACVPTDNPQEVLENTYE 61
 DB 31 EKLWTVYGVVPVWKEATTLFCASDRKAYDETEVHNWATHACVPTDNPQEVLEKNVTE 90

QY 62 HFNWKNMVEQMOEDIIISLDQSLKPCVKLTPLCVTLNCKDV-----NATNTNDSEG 115
 DB 91 NFNWKNMVEQMHEDIISLDQSLKPCVKLTPLCVTLNCTDLNATNGNDNTSSRG 150

QY 116 TMRGEIKNCSFNITTSIRDEVQKEYALFYKLDVVPIDNN-NTSVRLISCDSVTITQACP 174
 DB 151 MVGGEKNCNCFNITTIIRGKQKEYALFYKLDIAPIDNNSNNRYRLISCNTSVITQACP 210

QY 175 KISPEPIPIHYCAPAGFAILKNDKTFNGKPCKNVSTVQCTGIRPVVSTQLLNGSLA 234
 DB 211 KVSPEPIPIHYCAPAGFAILKCKDKFKNGKPCNTVSTVQCTGIRPVVSTQLLNGSLA 270

QY 235 EEEVIRSDNFTNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPRAFYTTGIIIGDIR 294
 DB 271 EEEVIRSANFADNAKVIIVQLNESVEINCTRPNNNTRKSIHIGPRAFYTTGIIIGDIR 330

QY 295 QAHCNISRAKNDTLKQIVIKLREQFNKTIIVFNHSSGGDPEIVMHSFNCSEGFYCNST 354
 DB 331 QAHCNLSRAKNDTLKIVIKLREQFNKTIIVFKHSSGGDPEIVTHSPNCSEGFYCNST 390

QY 355 QLFNSTWNTVEESNNTVENNTITLPCRKIQIINMWQEVGRAMYAPPPIRGQIRCSNITGL 414
 DB 391 QLFNSTWNTVEESNNTVENNTITLPCRKIQIINMWQEVGRAMYAPPPIRGQIRCSNITGL 450

QY 415 LLTRDGGINENGTEIFRPGGDMRDNMRSEFYKVKVIEPLGVAPTCKKRVVQREKRA 474
 DB 451 LLTRDGGPEDNKTEVFRPGGDMRDNMRSELYKYKVKIEPLGVAPTCKKRVVQREKRA 510

QY 475 VGIGAVFLGFLGAAGSTWGAASMTLTVOARLLLSGIVQOQNLLRAIEAQOQLLTWYG 534
 DB 511 VGIGAVFLGFLGAAGSTWGAASMTLTVOARLLLSGIVQOQNLLRAIEAQOQLLTWYG 570

QY 535 IKOLQARVLAVERYLGDQQLLGIWCSGKLTCTTAVPWNASWSNKSLLDRINWNTWMEWE 594
 DB 571 IKOLQARVLAVERYLGDQQLLGIWCSGKLTCTTAVPWNASWSNKSLLDRINWNTWMEWD 630

QY 595 REIDNNTSEIYTLIESQOQEKNEQELLELDKWASLWNVFDITWLMWY 643
 DB 631 REINNTYIILSLIESQOQEKNEQELLELDKWASLWNVFDITWLMWY 679

ALT 6
 AAW88113
 ID AAW88113 standard; Protein; 855 AA.
 XX AAW88113;
 AC AAW88113;
 XX
 DT 09-APR-1999 (first entry)
 XX
 DE Env protein of the BA-L strain of Human immunodeficiency virus type 1.
 XX
 KW HIV-1; HIV-1 strain BA-L; env protein; vaccine;
 KW immunotherapy; HIV infection; immunogen; HIV-1 diagnosis.
 XX
 OS Human immunodeficiency virus type 1.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 11 /note= "Gln encoded by CG"
 FT
 XX US5869313-A.
 XX
 PN 09-FEB-1999.
 XX
 PD 14-MAY-1996; 96US-0647714.
 XX
 PF

XX 17-OCT-1990; 90US-0599491.
 PR 25-FEB-1993; 93US-0022835.
 PR 15-FEB-1995; 95US-0388809.
 PR 14-MAY-1996; 96US-0647714.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 XX Franchini G, Gallo RC, Gartner S, Lori FC, Markham PD;
 PI Popovic M, Reitz MS;
 XX
 DR WPI; 1999-152779/13.
 DR N-PSDB; AAX04767.
 XX
 PT DNA encoding env protein of the human immune deficiency virus
 PT isolate BA-L - useful for producing protein for use in vaccines, as
 PT assay reagent and to generate antibodies
 XX
 XX Example 1; Fig 9A-C; 87pp; English.
 PS
 CC The present sequence represents the envelope protein of the BA-L
 CC (ATCC 40850) strain of Human immunodeficiency virus type 1 (HIV-1)
 CC strain MN-STL. BA-L is more typical of United States isolates of
 CC HIV-1 than previously known strains. Recombinant, complete env protein
 CC of the BA-L strain is used as a vaccine component and for immunotherapy
 CC of existing HIV infections, to detect HIV-specific antibodies, e.g. in
 CC donated blood, and as an immunogen to raise specific antibodies, for
 CC HIV-1 diagnosis.
 CC
 SQ Sequence 855 AA;

Query Match 88.7%; Score 3069.5; DB 20; Length 855;
 Best Local Similarity 88.1%; Pred. No. 3.5e-151;
 Matches 572; Conservative 27; Mismatches 43; Indels 7; Gaps 2;

QY 2 EKLWTVYGVVPVWKEATTLFCASDAKAYDETEVHNWATHACVPTDNPQEVLENTYE 61
 DB 31 EKLWTVYGVVPVWKEATTLFCASDRKAYDETEVHNWATHACVPTDNPQEVLEKNVTE 90

QY 62 HFNWKNMVEQMOEDIIISLDQSLKPCVKLTPLCVTLNCKDV-----NATNTNDSEG 115
 DB 91 NFNWKNMVEQMHEDIISLDQSLKPCVKLTPLCVTLNCTDLNATNGNDNTSSRG 150

QY 116 TMRGEIKNCSFNITTSIRDEVQKEYALFYKLDVVPIDNN-NTSVRLISCDSVTITQACP 174
 DB 151 MVGGEKNCNCFNITTIIRGKQKEYALFYKLDIAPIDNNSNNRYRLISCNTSVITQACP 210

QY 175 KISPEPIPIHYCAPAGFAILKNDKTFNGKPCKNVSTVQCTGIRPVVSTQLLNGSLA 234
 DB 211 KVSPEPIPIHYCAPAGFAILKCKDKFKNGKPCNTVSTVQCTGIRPVVSTQLLNGSLA 270

QY 235 EEEVIRSDNFTNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPRAFYTTGIIIGDIR 294
 DB 271 EEEVIRSANFADNAKVIIVQLNESVEINCTRPNNNTRKSIHIGPRAFYTTGIIIGDIR 330

QY 295 QAHCNISRAKNDTLKQIVIKLREQFNKTIIVFNHSSGGDPEIVMHSFNCSEGFYCNST 354
 DB 331 QAHCNLSRAKNDTLKIVIKLREQFNKTIIVFKHSSGGDPEIVTHSPNCSEGFYCNST 390

QY 355 QLFNSTWNTVEESNNTVENNTITLPCRKIQIINMWQEVGRAMYAPPPIRGQIRCSNITGL 414
 DB 391 QLFNSTWNTVEESNNTVENNTITLPCRKIQIINMWQEVGRAMYAPPPIRGQIRCSNITGL 450

QY 415 LLTRDGGINENGTEIFRPGGDMRDNMRSEFYKVKVIEPLGVAPTCKKRVVQREKRA 474
 DB 451 LLTRDGGPEDNKTEVFRPGGDMRDNMRSELYKYKVKIEPLGVAPTCKKRVVQREKRA 510

QY 475 VGIGAVFLGFLGAAGSTWGAASMTLTVOARLLLSGIVQOQNLLRAIEAQOQLLTWYG 534
 DB 511 VGIGAVFLGFLGAAGSTWGAASMTLTVOARLLLSGIVQOQNLLRAIEAQOQLLTWYG 570

QY 535 IKOLQARVLAVERYLGDQQLLGIWCSGKLTCTTAVPWNASWSNKSLLDRINWNTWMEWE 594
 DB 571 IKOLQARVLAVERYLGDQQLLGIWCSGKLTCTTAVPWNASWSNKSLLDRINWNTWMEWD 594

Qy	181	IDIHYCAPAGFAILKNDCKDTNGGPKCKNVSTVQCTHGIRPVVSTOLLNGSLAEBEWVI	246
Db	117	IDIHYCAPAGFAILKNDCKDTNGGPKCKNVSTVQCTHGIRPVVSTOLLNGSLAEBEWVI	176
Qy	241	RSDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPGRAFTVTGGEIIGDIRQAHGNI	236
Db	177	RSDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPGRAFTVTGGEIIGDIRQAHGNI	236
Qy	301	SRAKWNDDLKQIVIKLRQFENKTIIVFNHSSGGDPEIVMHSFNCEGEFFYCNSDTQLFNST	360
Db	237	SRAKWNDDLKQIVIKLRQFENKTIIVFNHSSGGDPEIVMHSFNCEGEFFYCNSDTQLFNST	236
Qy	361	WNNTEGNSNTEGNTITILPCRIKOIINMWQEVGKAMYPPIRGQIRCSSNITGLLLTRDG	420
Db	297	WNNTEGNSNTEGNTITILPCRIKOIINMWQEVGKAMYPPIRGQIRCSSNITGLLLTRDG	356
Qy	421	GINENGTEIFRPGCGDMEDNRSPEFYKVKVKTPELGVAPTCKKRVVQREKAVGIGAV	480
Db	357	GINENGTEIFRPGCGDMEDNRSPEFYKVKVKTPELGVAPTCKKRVVQREKAVGIGAV	416
Qy	481	FLFGFGAAGSTMGAAASMTLTVQARLLLSGIVQOQNLLRAIEAQORMQLQTLVWGIKQLQA	540
Db	417	FLFGFGAAGSTMGAAASMTLTVQARLLLSGIVQOQNLLRAIEAQORMQLQTLVWGIKQLQA	476
Qy	541	RVLAVERYLGDQQLLIGWCGSKLICCTAVPWNASWNSKSLDRIWNNMTWMEWEREIDNY	600
Db	477	RVLAVERYLGDQQLLIGWCGSKLICCTAVPWNASWNSKSLDRIWNNMTWMEWEREIDNY	536
Qy	601	TSEIYVTLTEESQNOQEKNEQELLELDKWSLWNWFDITNMLWY	643
Db	537	TSEIYVTLTEESQNOQEKNEQELLELDKWSLWNWFDITNMLWY	579
RESULT 8			
AAAY97073			
ID	AAAY97073 standard; protein; 847 AA.		
XX	AAAY97073;		
AC	AAAY97073;		
XX	31-OCT-2000 (first entry)		
DT	Variant HIV-1 SF162 Env gp160.		
XX	HIV-1; vaccine; anti-HIV; immunogenic; modified; envelope; Env;		
KW	CD4 binding region; V1/V2 loop; bridging sheet.		
XX	Human immunodeficiency virus type 1 isolate SF162.		
OS	WO2000039303-A2.		
XX	06-JUL-2000.		
XX	30-DEC-1999; 99WO-US31272.		
PF	31-DEC-1998; 98US-0114495.		
XX	29-SEP-1999; 99US-0156670.		
PR	(CHIR) CHIRON CORP.		
XX	Barnett S, Hartog K, Martin E;		
PI	WPI; 2000-465745/40.		
DR	Novel modified HIV Env polypeptides useful as immunizing agents and for		
PT	preparing a vaccine to elicit an immune response against a broad range		
PT	of HIV subtypes		
XX	Claim 5; Page 115-117; 139pp; English.		
PS	Novel immunogenic modified human immunodeficiency virus (HIV) envelope		
XX	(Env) polypeptides having an amino acid deleted or replaced in the region		
CC	corresponding to residues 420-436 or 119-123 and 199-210 relative to		
CC	isolate HXB-2 are disclosed. The modified Env polypeptide is based on HIV		

Db 265 IRSEFTDNTAKTIIIVOLKESVEINCTRPNNTRKSIITIGPRAFYATGDIIGDIRQAHCN 324
 Qy 300 ISRAKNDTLKQIVIKLREOFENKTIIVFNHSSGGDPEIIVHSHFNCGEFFYCNSTOLFNS 359
 Db 325 ISGEKWNNTLKQIVTKLQAQFGNKTIVFKQSGGDPPEIIVHSHFNCGEFFYCNSTOLFNS 384
 Qy 360 TWNNTEGSNNTGNTITLPCRIKQIINNMQEVGKAMYPPIRGQIRCSNITGLLLTRD 419
 Db 385 TW-NTTIGPNTNG-TITLPCRIKQIINRMQEVGKAMYPPIRGQIRCSNITGLLLTRD 442
 Qy 420 GGIN-ENGTEIPRPGGDMRDNRSSEFYKYVVKIEPLGVAPTCKRRVVQREKRAVGIG 478
 Db 443 GKEISNTTTEIPRPGGDMRDNRSSELYKYVVKTEPLGVAPTCKRRVVQREKRAVTLG 502
 Qy 479 AVFLGFLGAAGSTMGAASMTLTVOARLLLSGIVQQQNNLLRAEAQOHLQLTVMGIKQL 538
 Db 503 AMFLGFLGAAGSTMGARSLTLTVOARQLLSGIVQQQNNLLRAEAQOHLQLTVMGIKQL 562
 539 QARVLAVERYLGQOQLLGWCGSKLICCTTAVPWNASWSNKSLSLDQIWNNTWMEWEREID 598
 563 QARVLAVERYLKDQQLLGWCGSKLICCTTAVPWNASWSNKSLSLDQIWNNTWMEWEREID 622
 599 NYTSEIYTLIEESQOQKNEQELLELDKWSLWNNWFDITNMLWY 643
 623 NYTNLTITLIEESQOQKNEQELLELDKWSLWNNWFDISKMLWY 667

RESULT 10

ABB06211
 ID ABB06211 standard; Protein; 842 AA.

AC ABB06211;
 XX

DT 15-MAY-2002 (first entry)
 XX

DE HIV Env isolate SFI62 amino acid sequence.
 XX

KW Human immunodeficiency virus type C; antigenic HIV type C protein;
 KW immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
 KW immunostimulant; gene therapy.
 XX

OS Human immunodeficiency virus type C.
 OS Synthetic.
 XX

PN W0200204493-A2.
 XX

17-JAN-2002.
 XX

05-JUL-2001; 2001WO-US21241.
 XX

05-JUL-2000; 2000US-0610313.
 XX

(CHIR) CHIRON CORP.
 (UYST-) UNIV STELLENBOSCH.
 XX

Zur Megede J, Barnett SW, Engelbrecht S, Van Renaburg EJ;
 WPI; 2002-154920/20.
 XX

New polynucleotides encoding antigenic HIV Type C polypeptides, useful
 in applications including DNA immunization or generation of packaging
 cell lines, particularly in gene therapy -
 XX

Disclosure; Fig 105; 233pp; English.
 XX

The present invention describes expression cassettes comprising a
 polynucleotide sequence encoding a polypeptide comprising immunogenic
 HIV type C polypeptides. The expression cassettes comprise any of the
 HIV type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or
 Nef (i). (i) have immunostimulant activity and can be used in gene
 therapy. The HIV type C polynucleotides are useful in applications
 including DNA immunisation, generation of packaging cell lines, and

CC production of HIV Type C proteins. The polynucleotides are particularly
 CC useful in gene therapy and DNA immunisation applications. ABL39942 to
 CC ABL40054 and ABB06204 to ABB06215 represent sequences used in the
 CC exemplification of the present invention.

XX Sequence 842 AA;
 SQ

Query Match 87.4%; Score 3026; DB 23; Length 842;
 Best Local Similarity 88.2%; Pred. No. 6.2e-149;
 Matches 569; Conservative 29; Mismatches 43; Indels 4; Gaps 4;

Qy 1 VEKLVVTVYVYGVPMKEATTTILFCASDAKAYDTEVHNWATHACVPTDPNPQEVLENVT 60
 Db 25 VEKLVVTVYVYGVPMKEATTTILFCASDAKAYDTEVHNWATHACVPTDPNPQEVLENVT 84
 Qy 61 EHPNWKNNWVQMDEDIISLWDOSLKPCKVLTPCLVTINLCKDV-NATNTNDSEGTMER 119
 Db 85 ENFNWKNWVQMDEDIISLWDQSLKPCVKLTPLCTLHCTNLKLNKATKTSNWKEMDR 144
 Qy 120 GEIKNCSFNITTSIRDEVOKEYALFYKLDVDPIDNNNTSYRLISCDTSVITQACPISPE 179
 Db 145 GEIKNCSFKVTTISIRNKQKEYALFYKLDVDPIDNDNTSYKLINCNTSVTTQACPVSFE 204
 Qy 180 PPIHYCAPAGFAILLKNDKTFNGKPCXKXVSTVQCTHGIRPVVSTOLLNGLSLAEVV 239
 Db 205 PPIHYCAPAGFAILLKNDKFKNGSGPCTNVSTVQCTHGIRPVVSTOLLNGLSLAEVV 264
 Qy 240 IRSDFNTNAKTIIVQLKESVEINCTRPNNTRKSHIIGPRAFYTTGIIIGDIRQAHCN 299
 Db 265 IRSEFTDNTAKTIIIVOLKESVEINCTRPNNTRKSTTIGPRAFYATGDIIGDIRQAHCN 324
 Qy 300 ISRAKNDTLKQIVIKLREOFENKTIIVFNHSSGGDPEIIVHSHFNCGEFFYCNSTOLFNS 359
 Db 325 ISGEKWNNTLKQIVTKLQAQFGNKTIVFKQSGGDPPEIIVHSHFNCGEFFYCNSTOLFNS 384
 Qy 360 TWNNTEGSNNTGNTITLPCRIKQIINNMQEVGKAMYPPIRGQIRCSNITGLLLTRD 419
 Db 385 TW-NTTIGPNTNG-TITLPCRIKQIINRMQEVGKAMYPPIRGQIRCSNITGLLLTRD 442
 Qy 420 GGIN-ENGTEIPRPGGDMRDNRSSEFYKYVVKIEPLGVAPTCKRRVVQREKRAVGIG 478
 Db 443 GKEISNTTTEIPRPGGDMRDNRSSELYKYVVKTEPLGVAPTCKRRVVQREKRAVTLG 502
 Qy 479 AVFLGFLGAAGSTMGAASMTLTVOARLLLSGIVQQQNNLLRAEAQOHLQLTVMGIKQL 538
 Db 503 AMFLGFLGAAGSTMGARSLTLTVOARQLLSGIVQQQNNLLRAEAQOHLQLTVMGIKQL 562
 Qy 539 QARVLAVERYLGQOQLLGWCGSKLICCTTAVPWNASWSNKSLSLDQIWNNTWMEWEREID 598
 Db 563 QARVLAVERYLKDQQLLGWCGSKLICCTTAVPWNASWSNKSLSLDQIWNNTWMEWEREID 622
 Qy 599 NYTSEIYTLIEESQOQKNEQELLELDKWSLWNNWFDITNMLWY 643
 Db 623 NYTNLTITLIEESQOQKNEQELLELDKWSLWNNWFDISKMLWY 667

RESULT 11

AAR67725
 ID AAR67725 standard; Protein; 857 AA.

XX AAR67725;
 XX

DT 25-MAR-2003 (updated)
 DT 11-SEP-1995 (first entry)
 XX

DE gp120 from the HIV GNE16 isolate, clone 1.
 XX

KW HIV; human immunodeficiency virus; gp120; glycoprotein;
 KW GNE clone; GNE16 isolate.
 XX

OS Human immunodeficiency virus type 1.
 XX

PN W09428929-A1.

```

XX PD 22-DEC-1994.
XX PF 07-JUN-1994; 94WO-US06036.
XX PR 07-JUN-1993; 93US-0072833.
XX PA (GETH ) GENENTECH INC.
XX PI Berman PW, Nakamura GR;
XX PI WPI; 1995-036112/05.
XX DR N-PSDB; AAQ76019.
XX PT Use of HIV gp 120 polypeptide(s) - for developing probes for the
XX PT analysis, prevention and therapy of HIV infection
XX PS Claim 24; Page 32-35; 108pp; English.
XX

This protein is the gp120 (envelope glycoprotein) from the GNE16 isolate
of HIV-1 (human immunodeficiency virus type 1). There are neutralising
epitopes in the V2 and C4 domains of gp120, in addition to the
neutralising epitopes in the V3 domain. Although the amino acid sequences
of the neutralising epitopes in the V2, V4 and C4 domains are variable,
the amount of variation is highly constrained. This facilitates the
design of HIV subunit vaccines that can induce antibodies that neutralise
the most common HIV strains for a given geographic region. This invention
provides a multivalent gp120 subunit vaccine where the gp120 present in
the vaccine is from at least two HIV isolates which have different amino
acid sequences for a neutralising epitope in these regions.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX Sequence 857 AA;

Query Match 86.6%; Score 2999; DB 16; Length 857;
Best Local Similarity 85.4%; Pred. No. 1.6e-147;
Matches 556; Conservative 41; Mismatches 44; Indels 10; Gaps 3;

QY 3 KLVVTVYGVVPVKKEATTLFCASDAKAYDTEVHNWATHACVPTDPNPQEVVLENVTEH 62
DB 32 KLVVTVYGVVPVKETITTLFCASDAKAYDTEVHNWATHACVPTDPNPQEVVLENVTEH 91
QY 63 FNMKNMNVQEQHEDIISLWQSLKPCVKLTPLCVTLNCKDV-NATNTTNDSEGMERGE 121
DB 92 FNMKNMNVQEQHEDIISLWQSLKPCVKLTPLCVTLNCKDV-NATNTTNDSEGMERGE 151
QY 122 IKNCFSNITTSRDEVOKEYALFYKLDVVPIDN-----NTSVRLISCDSVTIQACP 174
DB 152 IKNCFSNITTSRDEVOKEYALFYKLDVVPIDN-----NTSVRLISCDSVTIQACP 211
QY 175 KLSFEPPIHYCAPAGFAILKCNKDTGKNGKPCXKSTVQCTHGIRPVVSTOLLNGSLA 234
DB 212 KVSFEPPIHYCAPAGFAILKCNKDTGKNGKPCXKSTVQCTHGIRPVVSTOLLNGSLA 271
QY 235 EEEVIRSDNFTNNAKTIIVQLKESVEINCTRPNNTRKSIHIGPCAFYTTGIIIGDIR 294
DB 272 EGEVIRSDNFTNNAKTIIVQLKESVEINCTRPNNTRKSIHIGPCAFYTTGIIIGDIR 331
QY 295 QAHCHISRAKNDTLKQIVIKLREQENKTIIVFNHSSGGDPPIVHSHFNCGEFFYCNS 354
DB 332 QAHCHLSRDMNTLQIVIKLREQENKTIIVFNHSSGGDPPIVHSHFNCGEFFYCNS 391
QY 355 QLFNSTWNN--TEGNNTEGNTITLPCRIKQIINWQEVGKAMYAPPIRGQIRCSSNIT 412
DB 392 QLFNSTWNTKVSNGTSTTEENSTITLPCRIKQIINWQEVGKAMYAPPIRGQIRCSSNIT 451
QY 413 GLLLTRDGGINGENGTETIFRPGGDMRDNRSEFYKVKVIEPLGVAPTKCRVRVQREK 472
DB 452 GLLLTRDGGSSNMSNETFRPGGDMRDNRSELYKVKVIEPLGVAPTKCRVRVQREK 511
QY 473 RAVGIGAVPLGLGAGSTMGAASTLTVQARLLSGIVQOQNLLRAIEAQOQLQTV 532
DB 512 RAVGIGAVPLGLGAGSTMGAASTLTVQARLLSGIVQOQNLLRAIEAQOHLQTV 571

QY 533 WGIKQARVLAVERYLGDQQLLGIWCGSKLICCTAVPWNASWNSKSLDRIWNNMTWME 592
DB 572 WGIKQARVLAVERYLGDQQLLGIWCGSKLICCTAVPWNASWNSKSLDRIWNNMTWME 631
QY 593 WEREIDNVTYSIYTLIEESQOQEKNEQELLELDKQASLWNNFDTITNMLWY 643
DB 632 WEREIENYTSIYTLIEESQOQEKNEQELLELDQWASLWNNFDTITNMLWY 682

RESULT 12
AAP80967
ID AAP80967 standard; protein; 851 AA.
XX AC AAP80967;
XX XX
XX 25-MAR-2003 (updated)
XX 15-NOV-1990 (first entry)
XX HIV protein HT6.
XX HIV; HT6; gp160; envelope protein; RF; AIDS.
XX Human immunodeficiency virus variant RF.
XX EP272858-A.
XX 29-JUN-1988.
XX 14-DEC-1987; 87EP-0310967.
XX 31-AUG-1987; 87US-0091481.
XX 15-DEC-1986; 86US-0941111.
XX (REPK ) REPLIGEN CORP.
XX Rusche J, Lynn D, Carson H, Putney S, Jellis CL;
XX WPI; 1988-176944/26.
XX N-PSDB; AAN80948.
XX Prodn. of recombinant HIV envelope proteins in insect cells -
XX useful as vaccine against AIDS and for diagnosis and therapy.
XX Disclosure; Page ?; 4pp; English.
XX The sequence is the result of cloning a hybrid envelope gene from
XX HIV variants BH10 and RF. A central portion of the RF gene was
XX used, the rest being from te distantly related variant BH10. The
XX resulting clone, pACHT6 produces a hybrid gp 160 envelope protein
XX with novel immunological and antigenic characteristics. It may
XX be used to as a vaccine and for diagnosis and therapy of AIDS.
XX See also AAP80966.
XX (Updated on 25-MAR-2003 to correct PA field.)
XX Sequence 851 AA;

Query Match 86.4%; Score 2991.5; DB 9; Length 851;
Best Local Similarity 86.2%; Pred. No. 3.8e-147;
Matches 557; Conservative 40; Mismatches 44; Indels 5; Gaps 4;

QY 2 EKLWTVYGVVPVKKEATTLTLCASDAKAYDTEVHNWATHACVPTDPNPQEVVLENVTE 61
DB 32 EKLWTVYGVVPVKKEATTLTLCASDAKAYDTEVHNWATHACVPTDPNPQEVVLENVTE 91
QY 62 HFNWKNMNVQEQHEDIISLWQSLKPCVKLTPLCVTLNCKDVNATNTTNDSEG--TWER 119
DB 92 HFNWKNMNVQEQHEDIISLWQSLKPCVKLTPLCVSLKCTDLKNDTNTNSSSGRMIMEX 151
QY 120 GEIKNCFSNITTSRDEVOKEYALFYKLDVVPIDNNTTSYRLISCDSVTIQACPISFE 179
DB 152 GEIKNCFSNMTSIRGKQVKEYAFYKLDIIPIDNDTTSYRLTSCNTSVITQACPISFE 211

```

180 PIPHYCAPAGFAILLKNDKTFNGKGPCKXVSTVQCTHGIRPVVSTQLLGSLABEEV 239
 212 PIPHYCAPAGFAILLKCNKTFNGCTVSTVQCTHGIRPVVSTQLLGSLABEEV 271
 240 IRSDFNNAKTIIVQLKESVEINCTPNNTNTRKSIHIGPRAPYTTGIIIGDIRAHCN 299
 272 IRSDFNNAKTIIVQLNASVQINCTPNNTNTRKSIHIGPRAPYTTGIIIGDIRAHCN 331
 300 ISRAKNDTLKQIVIKLEOFENKTIIVFNHSSGDDPEIVMHSFNCGEFEYCNSTOLFS 359
 332 LSAQWNTLQVVKLRQFDNKTIIVFTSSGGDDPEIVLHSGGGEFFYCNSTOLFS 391
 360 TNNNTSGSNNTGCTN-TITLPCRIKQIINNMQEVGRKAMYPPIRGQIRCSSNITGLLTR 418
 392 TW-NSTEGSNNTGNDTITLPCRIKQIIVNMQEVGRKAMYPPISSGQIKCISNITGLLTR 450
 419 DGGIN-ENGTEIFRPGGDMRDNRSEFYKYVVKIEPLGVAPTKCKRRVVOREKRAVGI 477
 451 DGGEDTNTTEIFRPGGDMRDNRSEFYKYVVKIEPLGVAPTKCKRRVVOREKRAVGI 510
 478 GAVFLGFLGAGSTMGASMTLTVOARLLLSGIVQOQNNLLRAIEAQOQMLQITVNGIKQ 537
 511 GAVFLGFLGAGSTMGASMTLTVOARLLLSGIVQOQNNLLRAIEAQOQMLQITVNGIKQ 570
 538 LQARLAVERYLGDQQLLGIWCGSGKLICTTAVPMNASWNSKSLDRINNMNTWWEVEREI 597
 571 LQARLAVERYLKDQQLLGIWCGSGKLICTTAVPMNASWNSKSLDRINNMNTWWEVEREI 630
 598 DNTSEIYTLIESQOQKNEQELLELDKWSLNMWFDITNWLWY 643
 631 NNTSLIHSIESQOQKNEQELLELDKWSLNMWFDITNWLWY 676

RESULT 13
 AAP60422
 ID AAP60422 standard; Protein; 868 AA.
 XX
 AC AAP60422;
 XX
 DT 25-MAR-2003 (updated)
 DT 20-AUG-1991 (first entry)
 XX
 DE Sequence of LAV virus env protein.
 XX
 KW AIDS vaccine; diagnosis; immunoassay; HIV; HTLV-III.
 XX
 OS Lymphadenopathy virus.
 XX
 WO8602383-A.
 PD 24-APR-1986.
 XX
 PF 18-OCT-1985; 85WO-EP00548.
 XX
 PR 18-OCT-1984; 84GB-0016013.
 PR 16-NOV-1984; 84GB-0029099.
 PR 21-JAN-1985; 85GB-0001473.
 XX
 (INSP) INST PASTEUR.
 PA (CNRS) CENT NAT RECH SCIENTIF.
 XX
 Montagnier L, Krust B, Charnaret F, Chermann JC, Barresinou F;
 PI Alizon M, Sonigo P;
 XX
 DR WPI; 1986-119166/18.
 DR N-PSDB; AAN60365.
 XX
 Purified glycoprotein and peptide(s) - are recognised by sera contg.
 PT antibodies against lymphadenopathy virus and useful in detecting
 PT AIDS antibodies or in vaccines
 XX
 Disclosure; Fig 4; 75pp; English.
 PS
 PS

CC The inventors claim a polypeptide which is recognised by sera of
 CC human origin contg. antibodies against the virus of
 CC lymphadenopathies (LAV) or acquired immune deficiency syndrome
 CC (AIDS). Also claimed are various peptides corresp. to the AA
 CC sequences deducible from proteins encoded by LAV DNA, defined by
 CC specific residues (e.g. 12-32, 37-46, 49-79, 88-153) in accordance
 CC with a formula given in the specification.
 CC (Updated on 25-MAR-2003 to correct PR field.)
 CC (Updated on 25-MAR-2003 to correct PA field.)
 CC (Updated on 25-MAR-2003 to correct PI field.)
 XX
 SQ Sequence 868 AA;
 Query Match 86.4%; Score 2989.5; DB 7; Length 868;
 Best Local Similarity 86.6%; Pred. No. 4.9e-147; Indels 15; Gaps 8;
 Matches 568; Conservative 33; Mismatches 40;
 QY 2 EKLWTVVYGVVPMKEATTTLCASDAKAYDEVHNVMAHACVPTDPNPQEVVLNVTE 61
 DB 39 EKLWTVVYGVVPMKEATTTLCASDAKAYDEVHNVMAHACVPTDPNPQEVVLNVTE 98
 QY 62 HFMKNNMVEOMQEDIIISLDQSLKPCVKLTPLCVTLNCKDV-NATNT-----TNDSG- 115
 DB 99 NFMKNNMVEOMQEDIIISLDQSLKPCVKLTPLCVTLNCKDV-NATNT-----TNDSG- 158
 QY 116 -TWERGEIKNCSEFNITTSIRDEVQKEYALFYKLDVVPIIDNNNTSYRLISCDTSVITQACP 174
 DB 159 MMEKEIKNCSEFNITTSIRDEVQKEYALFYKLDVVPIIDNNNTSYRLISCDTSVITQACP 218
 QY 175 KISFEPIPIHYCAPAGFAILLKNDKTFNGKGPCKXVSTVQCTHGIRPVVSTQLLGSLA 234
 DB 219 KVSFEPIPIHYCAPAGFAILLKNDKTFNGKGPCKXVSTVQCTHGIRPVVSTQLLGSLA 278
 QY 235 EEEVIRSDNFTNNAKTIIVQLKESVEINCTPNNTNTRKSIHIGPRAPYTTGIIIGDIRAHCN 292
 DB 279 EEEVIRSDNFTNNAKTIIVQLKESVEINCTPNNTNTRKSIHIGPRAPYTTGIIIGDIRAHCN 337
 QY 293 IRQAHCHNISRAKNDTLKQIVIKLEOFENKTIIVFNHSSGDDPEIVMHSFNCGEFEYCN 351
 DB 338 MRQAHCHNISRAKNDTLKQIVIKLEOFENKTIIVFNHSSGDDPEIVMHSFNCGEFEYCN 397
 QY 352 NSTQLFNSNTWNN---NTEGSNNTGCTN-TITLPCRIKQIINNMQEVGRKAMYPPIRGQIR 407
 DB 398 NSTQLFNSNTWNN---NTEGSNNTGCTN-TITLPCRIKQIINNMQEVGRKAMYPPIRGQIR 457
 QY 408 SSNTGLLITRDGGINENCTEIRPGGDMRDNRSEFYKYVVKIEPLGVAPTKCKRRV 467
 DB 458 SSNTGLLITRDGGINENCTEIRPGGDMRDNRSEFYKYVVKIEPLGVAPTKCKRRV 517
 QY 468 VOREKRAVGIAGVFLGFLGAGSTMGASMTLTVOARLLLSGIVQOQNNLLRAIEAQOQML 527
 DB 518 VOREKRAVGIAGVFLGFLGAGSTMGASMTLTVOARLLLSGIVQOQNNLLRAIEAQOQML 577
 QY 528 LQITVWGIKQLOARLAVERYLGDQQLLGIWCGSGKLICTTAVPMNASWNSKSLDRINNM 587
 DB 578 LQITVWGIKQLOARLAVERYLGDQQLLGIWCGSGKLICTTAVPMNASWNSKSLDRINNM 637
 QY 588 MTWMEWEREIDNVTSEIYTLIESQOQKNEQELLELDKWSLNMWFDITNWLWY 643
 DB 638 MTWMEWEREIDNVTSEIYTLIESQOQKNEQELLELDKWSLNMWFDITNWLWY 693
 RESULT 14
 AAO19389
 ID AAO19389 standard; Protein; 868 AA.
 XX
 AC AAO19389;
 XX
 DT 10-DEC-2002 (first entry)
 XX
 DE Lymphadenopathy-associated virus env protein.
 XX
 KW Lymphadenopathy-associated virus; LAV; HIV; human immunodeficiency virus;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 12:28:04 ; Search time 14.6215 Seconds
(without alignments)
1860.670 Million cell updates/sec

Title: US-10-032-162-13

Percent score: 3462
Sequence: 1 VEKLMVTYVGVVWPKAATT.....ELDKWASIMWPDITNIMLWY 643

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

1 number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3091.5	89.3	850	2	US-08-448-603A-28
2	3091.5	89.3	850	3	US-09-134-075-28
3	3091.5	89.3	850	4	US-09-492-739-28
4	3069.5	88.7	855	1	US-08-023-835-6
5	3069.5	88.7	855	1	US-08-388-809-6
6	3069.5	88.7	855	2	US-08-647-714-6
7	3027.5	87.4	855	3	US-07-956-483-14
8	2999	86.6	857	2	US-08-448-603A-30
9	2999	86.6	857	4	US-09-134-075-30
10	2999	86.6	857	4	US-09-492-739-30
11	2992.5	86.4	861	1	US-08-127-489A-14
12	2992.5	86.4	861	1	US-08-482-847-14
13	2992.5	86.4	861	3	US-07-956-483-10
14	2992.5	86.4	861	3	US-08-472-240A-7
15	2992.5	86.4	861	3	US-08-472-240A-7
16	2992.5	86.4	861	4	US-08-817-441-103
17	2984.5	86.2	861	3	US-07-956-483-16
18	2978	86.0	856	1	US-08-022-835-2
19	2978	86.0	856	1	US-08-388-809-2
20	2978	86.0	856	2	US-08-647-714-2
21	2978	86.0	856	3	US-07-956-483-11
22	2978	86.0	856	3	US-09-124-900-9
23	2978	86.0	863	1	US-08-463-210-11
24	2977	86.0	826	1	US-08-375-510-2
25	2977	86.0	826	2	US-08-487-657-2
26	2977	86.0	854	4	US-09-309-572-23
27	2969	85.8	857	1	US-08-022-835-4

28	2969	85.8	857	1	US-08-388-809-4	Sequence 4, Appli
29	2969	85.8	857	2	US-08-647-714-4	Sequence 4, Appli
30	2966	85.7	887	3	US-08-472-240A-5	Sequence 5, Appli
31	2964	85.6	855	3	US-07-956-483-15	Sequence 15, Appli
32	2964	85.6	856	2	US-07-916-098A-2	Sequence 2, Appli
33	2963.5	85.6	839	3	US-08-472-240A-10	Sequence 10, Appli
34	2953	85.3	856	4	US-09-337-387-11	Sequence 11, Appli
35	2947.5	85.1	880	2	US-08-788-815-7	Sequence 7, Appli
36	2947.5	85.1	880	3	US-09-157-963-7	Sequence 11, Appli
37	2935	84.8	856	1	US-08-375-100-1	Sequence 1, Appli
38	2923	84.4	887	3	US-08-472-240A-6	Sequence 6, Appli
39	2917	84.3	865	3	US-07-956-483-13	Sequence 13, Appli
40	2910.5	84.1	726	4	US-09-337-387-3	Sequence 3, Appli
41	2907	84.0	887	3	US-08-472-240A-2	Sequence 2, Appli
42	2897.5	83.7	759	4	US-09-337-387-12	Sequence 12, Appli
43	2891.5	83.5	657	4	US-09-256-194-2	Sequence 2, Appli
44	2888.5	83.4	665	2	US-08-448-603A-32	Sequence 32, Appli
45	2888.5	83.4	665	3	US-09-134-075-32	Sequence 32, Appli

ALIGNMENTS

```
RESULT 1
US-08-448-603A-28
; Sequence 28, Application US/08448603A
; Patent No. 5864027
;
; GENERAL INFORMATION:
; APPLICANT: Berkman, Phillip W.
; APPLICANT: Nakamura, Gerald R.
; TITLE OF INVENTION: HIV Envelope Polypeptides
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESS: McCutchen, Doyle, Brown & Emersen, LLP
; STREET: 3 Embarcadero Center
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,603A
; FILING DATE: 07-JUN-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/072,833
; FILING DATE: 07-JUN-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Halliday, Emily
; REGISTRATION NUMBER: 38903
; REFERENCE/DOCKET NUMBER: 14918-704
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-393-2000
; TELEFAX: 415-393-2286
;
; TELE:
;
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 850 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; US-08-448-603A-28
;
; Query Match 89.3%; Score 3091.5; DB 2; Length 850;
; Best Local Similarity 89.7%; Pred. No. 3e+250;
; Matches 581; Conservative 26; Mismatches 32; Indels 9; Gaps 5;
;
; 2 EKLMTVYVGVVWPKAATTLLFCASDAKAYDEVHVNWATHACVPTDPNPEVLENTVE 61
; |||||
```



```

Db      31 EKLMTVYVGVPMKEATTTLLFCASDAKAYDTEVHVMATHACVPTDPNPOEIGLENVTE 90
Qy      62 HNNMKNMVEQMEQEDIIISLMDOSLKPVCVLTPLCVTLNCKDV-NANTNTNDSBGMERG 120
Db      91 NFNMMKNMVEQMEQEDIIISLMDOSLKPVCVLTPLCVTLNCKDV-NANTNTNDSBGMERG 150
Qy      121 EIKNCSFNITTSIRDEYQKEVALFYKLDVYPIDNNNTSYRLISCDTSVITQACPISFEP 180
Db      151 EIKNCSFNVTTSIRDKMKEVALFYKLDVYPIDNDNTSYRLISCNISVITQACPISFEP 210
Qy      181 IPIHYCAPAGFALLKCKDKTFNGKPCKNVSTVQCTHGIRPVVSTOLLNGSLAEEVVI 240
Db      211 IPIHYCAPAGFALLKCKDKTFNGKPCKNVSTVQCTHGIRPVVSTOLLNGSLAEEVVI 270
Qy      241 RSDNFTNNAITIIYOLKESVEINCTRPNNTRKSHIGPRAPYATGEIIGDIRQAHCNL 300
Db      271 RSNFNSNATIIIVOLNESVEINCTRPNNTRSHIGPRAPYATGEIIGDIRQAHCNL 330
Qy      301 SRAMKNDTLKQIYIKLREOFENKTIIVFNHSSGDPPEIVMHSFNCGEFFYCNSDTLPNST 360
Db      331 SSTRKMNNTLKQIYIKLREHF-NKTIIVFNHSSGDPPEIVMHSFNCGEFFYCNSDTLPNST 389
Qy      361 WN-----NTEGSNNTGNTITLPCRIKQIINMVEGKAMVAPIRGQIRCSNITGL 415
Db      390 WNYTYTNNTGSGNDT-GRNITLQCRIKQIINMVEGKAMVAPIRGQIRCSNITGL 448
Qy      416 LTRDGINENGTIFRPGGDMRDNMRSEFYKYVVKIEPLGVAPTYCKRKYVQREKRAV 475
Db      449 LTRDGG-NNSETEIFRPGGDMRDNMRSELYKYVVKIEPLGVAPTYCKRKYVQREKRAV 507
Qy      476 GIGAVFLGFLGAGSTMGASMTLTVQARLLSGIVQOQNLLRAIEAQOQRLQITWGI 535
Db      508 GIGAVFLGFLGAGSTMGASMTLTVQARLLSGIVQOQNLLRAIEAQOQRLQITWGI 567
Qy      536 KQLOARVLAVERYLKQOQLLGIWCGSGKLICTAVPNWASWSNKSJDRINNNTWEMER 595
Db      568 KQLOARVLAVERYLKQOQLLGIWCGSGKLICTAVPNWASWSNKSJDRINNNTWEMER 627
Qy      596 EIDNYSIETLTLIESQNOQEKNOELLEDKASLNMWPDITWMLY 643
Db      628 EIDNYSIETLTLIESQNOQEKNOELLEDKASLNMWPDITWMLY 675

RESULT 2
US-09-134-075-28
; Sequence 28, Application US/09134075
; Patent No. 6042836
; GENERAL INFORMATION:
; APPLICANT: Berman, Phillip W.
; APPLICANT: Nakamura, Gerald R.
; TITLE OF INVENTION: HIV Envelope Polypeptides
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McCutchen, Doyle, Brown & Eneersen, LLP
; STREET: 3 Embarcadero Center
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: PatsISO for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/134.075
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/448,603
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Haliday, Emily

```

```

; REGISTRATION NUMBER: 38903
; REFERENCE/DOCKET NUMBER: 14918-704
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-393-2000
; TELEFAX: 415-393-2286
;
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 850 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; US-09-134-075-28
;
; Query Match: 89.3%; Score 3091.5; DB 3; Length 850;
; Best Local Similarity 89.7%; Pred. No. 3e-250;
; Matches 581; Conservative 26; Mismatches 32; Indels 9; Gaps 5;
;
; 2 EKLMTVYVGVPMKEATTTLLFCASDAKAYDTEVHVMATHACVPTDPNPOEIGLENVTE 61
; 31 EKLMTVYVGVPMKEATTTLLFCASDAKAYDTEVHVMATHACVPTDPNPOEIGLENVTE 90
;
; 62 HNNMKNMVEQMEQEDIIISLMDOSLKPVCVLTPLCVTLNCKDV-NANTNTNDSBGMERG 120
; 91 NFNMMKNMVEQMEQEDIIISLMDOSLKPVCVLTPLCVTLNCKDV-NANTNTNDSBGMERG 150
;
; 121 EIKNCSFNITTSIRDEYQKEVALFYKLDVYPIDNNNTSYRLISCDTSVITQACPISFEP 180
; 151 EIKNCSFNVTTSIRDKMKEVALFYKLDVYPIDNDNTSYRLISCNISVITQACPISFEP 210
;
; 181 IPIHYCAPAGFALLKCKDKTFNGKPCKNVSTVQCTHGIRPVVSTOLLNGSLAEEVVI 240
; 211 IPIHYCAPAGFALLKCKDKTFNGKPCKNVSTVQCTHGIRPVVSTOLLNGSLAEEVVI 270
;
; 241 RSDNFTNNAITIIYOLKESVEINCTRPNNTRKSHIGPRAPYATGEIIGDIRQAHCNL 300
; 271 RSNFNSNATIIIVOLNESVEINCTRPNNTRSHIGPRAPYATGEIIGDIRQAHCNL 330
;
; 301 SRAMKNDTLKQIYIKLREOFENKTIIVFNHSSGDPPEIVMHSFNCGEFFYCNSDTLPNST 360
; 331 SSTRKMNNTLKQIYIKLREHF-NKTIIVFNHSSGDPPEIVMHSFNCGEFFYCNSDTLPNST 389
;
; 361 WN-----NTEGSNNTGNTITLPCRIKQIINMVEGKAMVAPIRGQIRCSNITGL 415
; 390 WNYTYTNNTGSGNDT-GRNITLQCRIKQIINMVEGKAMVAPIRGQIRCSNITGL 448
;
; 416 LTRDGINENGTIFRPGGDMRDNMRSEFYKYVVKIEPLGVAPTYCKRKYVQREKRAV 475
; 449 LTRDGG-NNSETEIFRPGGDMRDNMRSELYKYVVKIEPLGVAPTYCKRKYVQREKRAV 507
;
; 476 GIGAVFLGFLGAGSTMGASMTLTVQARLLSGIVQOQNLLRAIEAQOQRLQITWGI 535
; 508 GIGAVFLGFLGAGSTMGASMTLTVQARLLSGIVQOQNLLRAIEAQOQRLQITWGI 567
;
; 536 KQLOARVLAVERYLKQOQLLGIWCGSGKLICTAVPNWASWSNKSJDRINNNTWEMER 595
; 568 KQLOARVLAVERYLKQOQLLGIWCGSGKLICTAVPNWASWSNKSJDRINNNTWEMER 627
;
; 596 EIDNYSIETLTLIESQNOQEKNOELLEDKASLNMWPDITWMLY 643
; 628 EIDNYSIETLTLIESQNOQEKNOELLEDKASLNMWPDITWMLY 675

RESULT 3
US-09-492-739-28
; Sequence 28, Application US/09492739
; Patent No. 6331404
; GENERAL INFORMATION:
; APPLICANT: Berman, Phillip W.
; APPLICANT: Nakamura, Gerald R.
; TITLE OF INVENTION: HIV Envelope Polypeptides
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:

```


ADDRESSEE: McCutchen, Doyle, Brown & Enersen, LLP
 STREET: 3 Embarcadero Center
 CITY: San Francisco
 STATE: CA
 COUNTRY: USA
 ZIP: 94111

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: PASEQ for Windows Version 2.0.

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/492,739
 FILING DATE: 27-Jan-2000
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 09/134,075
 FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Haliday, Emily
 REGISTRATION NUMBER: 38903
 REFERENCE/DOCKET NUMBER: 14918-704

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-393-2000
 TELEFAX: 415-393-2286
 TELEX: <Unknown>
 INFORMATION FOR SEQ ID NO: 28:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 850 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 28:
 US-09-492-739-28

Query Match 89.3%; Score 3091.5; DB 4; Length 850;
 Best Local Similarity 89.7%; Pred. No. 3e-250;
 Matches 581; Conservative 26; Mismatches 32; Indels 9; Gaps 5;

2 EKLWTVVYGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTE 61
 31 EKLWTVVYGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTE 90
 62 HFNWKNMWMQEMODIISLMDOSLKPCKVLTPLCVTLNCKOV-NATNTNDSGEMRG 120
 91 NFNWKNMWMQEMODIISLMDOSLKPCKVLTPLCVTLNCKOV-NATNTNDSGEMRG 150
 121 EIKKNSFNTTISRDEVOKEYALFYKLDVVPIDNNNTSYRLISCTSVITQACPKISFEP 180
 151 EIKKNSFNTTISRDEVOKEYALFYKLDVVPIDNNNTSYRLISCTSVITQACPKISFEP 210
 181 IPIHCAAGAPAILKCDKXTENGKPCCKVSTVQCTHGIRPVSTOLLNGLAEVEYVI 240
 211 IPIHCAAGAPAILKCDKXTENGKPCCKVSTVQCTHGIRPVSTOLLNGLAEVEYVI 270
 241 RSDNFTNAKTIIVOLKESVEINCTRPNNNTRKSHIGRAFYTTGRIIGIRAHGNI 300
 271 RSDNFTNAKTIIVOLKESVEINCTRPNNNTRKSHIGRAFYTTGRIIGIRAHGNI 330
 301 SRAKNNDTLKOIVIKLREOFENKTIIVFNHSSGDEPEIWMHSGCEFFYCNSLTQFNST 360
 331 SRAKNNDTLKOIVIKLREOFENKTIIVFNHSSGDEPEIWMHSGCEFFYCNSLTQFNST 389
 361 WN-----NNTGCSNNTGNTTTLPCRIQIIMMOEVBGKAMAPAPIRQIIRCSSNITGL 415
 390 WNYTYTNNTRGSDNT-GRNITLQCRKQIIMMOEVBGKAMAPAPIRQIIRCSSNITGL 448
 416 LTRDGIENGTEIRPGGDMRDMRSEFYKYVVKLEPIGVAPTKCRVRVQREKXAV 475
 449 LTRDGIENGTEIRPGGDMRDMRSEFYKYVVKLEPIGVAPTKCRVRVQREKXAV 507
 476 GIGAVFLGAGSTMGASMTLTVQARLLLSGIVQOQNNLLRAIEAQRMQLQTVWGI 535

DB 508 GIGAVFLGAGSTMGASMTLTVQARLLLSGIVQOQNNLLRAIEAQRMQLQTVWGI 567
 QY 536 KOLARVLAVERYLGDQOLGIMGCSGLICTAVPMWASNSKSLDRINNTMTMEWER 595
 DB 568 KOLARVLAVERYLGDQOLGIMGCSGLICTAVPMWASNSKSLDRINNTMTMEWER 627
 QY 596 EIDNVTSEIYTLIESQNOQEKNEOELLELDKMASLWNPFDITWLVY 643
 DB 628 EIDNVTSEIYTLIESQNOQEKNEOELLELDKMASLWNPFDITWLVY 675

RESULT 4
 US-08-022-835-6
 Sequence 6, Application US/08022835
 Patent No. 5420030

GENERAL INFORMATION:
 APPLICANT: Reitz Jr., Marvin S.
 APPLICANT: Franchini, Genoveffa
 APPLICANT: Markham, Phillip D.
 APPLICANT: Gallo, Robert C.
 APPLICANT: Lori, Franco C.
 APPLICANT: Popovic, Mikulas
 APPLICANT: Garnier, Suzanne
 TITLE OF INVENTION: MOLECULAR CLONES OF HIV-1 AND USES
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
 STREET: Eleventh Floor, 1615 L. Street, N.W.
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20036-5601

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/022,835
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/599,491
 FILING DATE: 17-OCT-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Scott, Watson T.
 REGISTRATION NUMBER: 26,581
 REFERENCE/DOCKET NUMBER: WTS/5683/78507/SRL
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 861-3000
 TELEFAX: (202) 822-0944
 TELEX: 6714627 CUSH
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 855 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein

US-08-022-835-6

Query Match 88.7%; Score 3069.5; DB 1; Length 855;
 Best Local Similarity 88.1%; Pred. No. 2.1e-248;
 Matches 572; Conservative 27; Mismatches 43; Indels 7; Gaps 2;

2 EKLWTVVYGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTE 61
 31 EKLWTVVYGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTE 90
 62 HFNWKNMWMQEMODIISLMDOSLKPCKVLTPLCVTLNCKOV-----NATNTNDSG 115
 91 NFNWKNMWMQEMODIISLMDOSLKPCKVLTPLCVTLNCKOV-----NATNTNDSG 150

0Y 116 TMEBEIKNCSNITTSIDSEVOKEALPYKLDVNI-DNN-NSYSLISCDTSVTOACP 174
: : : : :
Db 151 MWGSEMKNCSNITTINIRKQKQEALYYKDIAIDNNSNRKYLSINTSVITOACP 210

0Y 175 KISPEPIPHYCAPAGFALKCNCKTENGSKPCKNSTVQCTHGIIPPVSTOLLNGSLA 234
:
Db 211 KVSFEPIPIHYCAPAGFALKCKDKFKFNKGPCPTNSTVQCTHGIIPPVSTOLLNGSLA 270

0Y 235 EEEVVIRSNPFNNMKTIIVOLKESEINCTRPNNTTRKS IHGPRAPFTTBETIIGDIR 294
: : : : :
Db 271 EEEVVIRSNPFADNAVIIIVOLNESVEINCTRPNNTTRKS IHGPRAPFTTBETIIGDIR 330

0Y 295 QAHCHNISAKNMDDLKQIIVIKLREOPENRTIYFNHSSGGDEIYMHSFNCGEFFYCNS 354
:
Db 331 QAHCULSRKAMNDTLNKTIYIKLRQEGCNTIYPKHSSGGDEIYVHSFNCGGFFYCNS 350

0Y 355 QLFNSTWNNNTSGSNTEGNITTLPCRIOIIMNOEVCAMYPPIRQIRCSSNITGL 414
:
Db 391 QLFNSTWNTYESNNTVENNITTLPCRIOIIMNOEVGRAMYAPPIRQIRCSSNITGL 450

0Y 415 LITRBGINENGETELFRPGGDMRDWRBSEFFYKKVKIEPLGVATPKCRARYOBEKRA 474
: : : : :
Db 451 LITRBGPEDNTEVFRRPGGDMRDWRBSELKYKKVKIEPLGVATPKARRVVOBEKRA 510

0Y 475 VGIGAVPFLGFALAAGSTMGAASMTLVQARILLSGIVOOONNLRLAI EAQRBLQLTWVG 534
:
Db 511 VGIQAVPFLGFALAAGSTMGAANTLVQARLLLSGIVOOONNLRLAI EAQHLLQLTWVG 570

0Y 535 IKOLARVLAVERYLGDOQLLGWGSCGKLICTTAVPWNASWNSKSLDRIMNMTWMWE 594
:
Db 571 IKOLARVLAVERYLRDQQLLGWGSCGKLICTTAVPWNASWNSKSLINKMDNMTIEMD 630

0Y 595 REINDYTSEIYLIEBSNQOEKNEDELLELDKWSLMNWPDITNMLWY 643
: : : : :
Db 631 REINNYTSIIYSILIESNQOEKNEDELLELDKWSLMNWPDITKMLWY 679

RESULT 5
US-08-388-809-6
Sequence 6, Application US/08388809
Patent No. 5576000
GENERAL INFORMATION:
APPLICANT: REITZ, JR., MARVIN S.; FRANCHINI,
APPLICANT: GENOVEFFA; MARKHAM, PHILIP D. GALLO, ROBERT
APPLICANT: C.; LORI, FRANCO C.; POPOVIC, MIKULAS; AND
APPLICANT: GARTNER, SUZANNE
TITLE OF INVENTION: MOLECULAR CLONES OF HIV-1
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & PINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK, 3.5"
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/388, 809
FILING DATE: 15-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/022, 835
FILING DATE: 25-FEB-1993
APPLICATION NUMBER: US 07/599, 491
FILING DATE: 17-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: LESLIE A. SERUNIAN
REGISTRATION NUMBER: 35,353
REFERENCE/DOCKET NUMBER: 2026-40920S2

```

? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (212) 758-4800
? TELEFAX: (212) 751-6849
? TELEX: 421792
? INFORMATION FOR SEQ ID NO: 6:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 855 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
US-08-368-809-6

```

Query Match	88.7%;	Score 3069.5;	DB 1;	Length 855;
Best Local Similarity	88.1%;	Pred. No. 2.1e-248;		
Matches 572;	Conservative 27;	Mismatches 43;	Indels 7;	Gaps 2.

OY	2	EKLWATVYGVPAWKEATTTTFCASDAAYOTEVHNWATACTVPTDNPQEVLEANTE	61
Db	31	EKLWATVYGVPAWKEATTTTFCASDRAYOTEVHNWATACTVPTDNPQEVLEKANTE	90
OY	62	HNPNKNNMVEQWQEDIIISLWDQSLKPCVKLTPLCVTLNCKDV'-----NATNTTNDSEG	115
Db	91	NPNMKNMNVQWQMEDITISLWDQSLKPCVKLTPLCVTLNCTDNLNATNGDNTNTTSSRG	150
OY	116	TMERGEIKNCSFNITTSIRDEVOKEYALFYLDVAVPIDNN'NTSVRLISCDTSVTAQCP	174
Db	151	MVGGEMKNCSFNITTNIRKVOKEYALFYLDIAPIDNNNNRRLISCTSVTAQCP	210
OY	175	KISEFPIPIHCAQAGFAILKCNKTFGKGPCKNVSIVQCTHGIRPVVSTOLLNGSLA	234
Db	211	KVSEFPIPIHCAQAGFAILKCKDKKFKGKPCPTNVSVQCTHGIRPVVSTOLLNGSLA	270
OY	235	EEVAVIRSDNFNTNAKTIIVOLKESVBINCTRPNNNTRKSIHIGBPAFYTTGELIGDIR	294
Db	271	EEVAVIRSANFADAKVIIVOLNESVBEINCTRPNNNTRKSIHIGBPAFYTTGELIGDIR	330
OY	295	QAHCNISAKKNDPLKOIVIKLREQENKTIIVPHNSGCGDEIIVHNSNCGEFPYCNST	354
Db	331	QAHCNLSAKKNDPLNKVIKLRQDFGKTIIVFKHSSGCGDEIIVHNSNCGEFPYCNST	390
OY	355	QLFNSVNNNTGSGNTEGNTITLPCRIKOIINMWQEVKAMYPAPIRGOIRCSSNITGL	414
Db	391	QLFNSVNNVTEESNNTVENNTITLPCRIKOIINMWQEVGRAMYAPPIGOIRCSSNITGL	450
OY	415	LLTTDDGGINNGTEIFRPGGDMDDMNSSEYKKVVKYIEPLGVAPTCKRXRVVQREKRA	474
Db	451	LLTTDDGGEDDKTEIFRPGGDMDDMNSSELYKKVVKYIEPLGVAPTAKRRVVQREKRA	510
OY	475	VGIGAVPLFGFAGAGSTGGAASMTLTVQARLLLSGIVQOQNNLLRAIEAQORMLQLTWYG	534
Db	511	VGIGAVPLFGFAGAGSTGGAASMTLTVQARLLLSGIVQOQNNLLRAIEAQOHLQLTWYG	570
OY	535	IKOLQARVLAVERYLQDOQLLIGWCCSKLICTPAVPNANASWKSXLDRIINNNTWMEWE	594
Db	571	IKOLQARVLAVERYLRDQQLLIGWCCSKLICTTAVPNNASWKSXKXLDRIINNNTWMEWD	630
OY	595	REINDYTSIETLLIEESONQOEKNOEELIEDKNASLNNWEDITWMLY	643
Db	631	REINDYTSIETLLIEESONQOEKNOEELIEDKNASLNNWEDITWMLY	679
RESULT 6			
US-08-647-714-6			
Sequence 6, Application US/08647714			
Patent No. 5869313			
GENERAL INFORMATION:			
APPLICANT: REITZ, JR., MARVIN S.; FRANCHINI,			
APPLICANT: GENOVEFA; MARGHAM, PHILIP D. GALLO, ROBERT			
APPLICANT: C.; LORI, FRANCO C.; POPOVIC, MIKULAS; AND			
APPLICANT: GARTNER, SUZANNE			
TITLE OF INVENTION: MOLECULAR CLONES OF HIV-1			
TITLE OF INVENTION: AND USES THEREOF			
NUMBER OF SEQUENCES: 6			

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK, 3.5"
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/647,714
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/388,809
FILING DATE: 15-FEB-1995
APPLICATION NUMBER: US 08/022,835
FILING DATE: 25-FEB-1993
APPLICATION NUMBER: US 07/599,491
FILING DATE: 17-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: LESLIE A. SERUNIAN
REGISTRATION NUMBER: 35,353
REFERENCE/DOCKET NUMBER: 2026-4092US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 855 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-647-714-6

Query Match      88.7%; Score 3069.5; DB 2; Length 855;
Best Local Similarity 88.1%; Pred. No. 2,1e-248;
Matches 572; Conservative 27; Mismatches 43; Indels 7; Gaps 2;

QY 2 EKLWTVYVYGVVWKEATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNENTE 61
DB 31 EKLWTVYVYGVVWKEATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNENTE 90
QY 62 HFNWKNWVEQWQEDIIISLWDSLKPCVKLTPLCVTLNCKDV-----NATNTTNDSEG 115
DB 91 NFNWKNWVEQWQEDIIISLWDSLKPCVKLTPLCVTLNCKDV-----NATNTTNDSEG 150
QY 116 TMRGEIKNCSFNTTTSIRDEVQKEYALFYKLDVVPIDNN-NTSYRLISCTSVITQCP 174
DB 151 WVGGEKNCSPNTTNRGRVQKEYALFYKLDVVPIDNN-NTSYRLISCTSVITQCP 210
QY 175 KISFEPIPIHCAFGAFAILKCNKDTFNGKPCPKVSTVOCTHGRPVYSTOLLNGSLA 234
DB 211 KVSFEPIPIHCAFGAFAILKCKDKKFKGKPCCTVSTVOCTHGRPVYSTOLLNGSLA 270
QY 235 EEEVVISRDNFTNNAKTIIIVQLKESVVEINCTRPNNNTKSIHIGGRAFYTTGELIGDIR 294
DB 271 EEEVVISRDNFTNNAKTIIIVQLKESVVEINCTRPNNNTKSIHIGGRAFYTTGELIGDIR 330
QY 295 QAHNINISAKKNDPLKQIVILKREOFENKTIIVFNHSSGDEBEIVHSHNGCEFFYCNST 354
DB 331 QAHNINISAKKNDPLKQIVILKREOFENKTIIVFNHSSGDEBEIVHSHNGCEFFYCNST 390
QY 355 QLFNSTWNNTEGSGNTTGTITLLPCRKKQIINNMQEVGKAMVYAPPIRGQIRCSNITGL 414
DB 391 QLFNSTWNNTEGSGNTTGTITLLPCRKKQIINNMQEVGKAMVYAPPIRGQIRCSNITGL 450
QY 415 LLTBDGGINENGTEIFRPDGGDMRDNMWSEFYKYRVKVIKIEPLGVAFTYCKKRVVQREKRA 474
DB 451 LLTBDGGINENGTEIFRPDGGDMRDNMWSEFYKYRVKVIKIEPLGVAFTYCKKRVVQREKRA 510

```

```

QY 475 VGIGAVFLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOQNNLRAIEAQOQMLQLTWVG 534
DB 511 VGIGAVFLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOQNNLRAIEAQOQMLQLTWVG 570
QY 535 IKOLQARVLAVERYLDQOQLIGWCSGKLICTAVPNNASMSNKLDRINNNNTWMEWE 594
DB 571 IKOLQARVLAVERYLDQOQLIGWCSGKLICTAVPNNASMSNKLDRINNNNTWMEWE 630
QY 595 REINDYTSIYTLIEESQOQEKNOEELIEDKASLNNWFDITNWLWY 643
DB 631 REINDYTSIYTLIEESQOQEKNOEELIEDKASLNNWFDITNWLWY 679

RESULT 7
US-07-956-483-14
Sequence 14, Application US/07956483
Patent No. 6261799
GENERAL INFORMATION:
APPLICANT: KIENY, Marie-Paule
TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED
NUMBER OF INVENTION: 9P160 VARIANT
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P. O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/956,483
FILING DATE: 31-DEC-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 92/19742
FILING DATE: 12-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 05392
FILING DATE: 02-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Crane-Feury, Sharon E
REGISTRATION NUMBER: 36,113
REFERENCE/DOCKET NUMBER: 017753-005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 855 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-07-956-483-14

Query Match      87.4%; Score 3027.5; DB 3; Length 855;
Best Local Similarity 87.2%; Pred. No. 6.9e-245;
Matches 571; Conservative 33; Mismatches 34; Indels 17; Gaps 7;

QY 2 EKLWTVYVYGVVWKEATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNENTE 61
DB 31 EKLWTVYVYGVVWKEATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNENTE 90
QY 62 HFNWKNWVEQWQEDIIISLWDSLKPCVKLTPLCVTLNCKDV-----NATNTTNDSEG 115
DB 91 NFNWKNWVEQWQEDIIISLWDSLKPCVKLTPLCVTLNCKDV-----NATNTTNDSEG 150

```

Qy	116	TMERGEJHONSFPITTSIDEXOKEALYKLDVPI	DINNNSYALSCDPSVITLOCPK	175
Db	151	KMEGERTNCSFPITTSIRSXOKEALYKLDVPID-	NTSYLLINCNSVITLOACP	208
Qy	176	ISPEPIPIHCADAPFAILKCDKDTENGKPCKN	SVYQCTHGIRPVVSTOLLNGSLAE	235
Db	209	VSEPIPIPIHCA-RWFAILINCNNKKENGSPCT	NSVSTVOCTHGIRPVVSTHLLNGSLAE	267
Qy	236	EEVYIRSDNTNNAKTIIVOLKESVINC	TRPNNTKRSIHIGPERAYTTEIGIRQ	299
Db	268	EEVYIRSDNTNNAKTIIVOLKESVINC	TRPNNTKRSIHIGPERAYATGDIIGIRQ	322
Qy	296	AHCNISRAKXNDTLKOIVIKLREOFENKTI	IVENHSGGPEIYVMSFNCGEFFYCNSTQ	355
Db	328	AHCNISRAKXNDTLKOIVIKLREQFNKTI	IIVRSGGPEIYVMSFNCGEFFYCNSTQ	387
Qy	356	LEFNSTNNNNNEGSNNTEGN-TTLPRIOI	INMMOEUKAMYPAPINGOIRGCSNITGL	414
Db	388	LEFSIT-NGTEGSNNNGDNTIILPRIKEL	IINMOEUKAMYPAPINGOVKCSNITGL	446
Qy	415	LLTRDGINENG-----TEIFRPGGDM	DMNRSEFYKXIKVLEPLGVAETPKRRVYQ	465
Db	447	LLTRDGENSGXGKNENTEIFRPGGDM	DMNRSELYKXIKVLEPLGVAETPKARRVYQ	508
Qy	470	REKRAV-ICAVPLGFGAAGSTMGASMTL	VOARLLSGIVOOONMLLAIRAQOML	528
Db	507	REKRAVGTIAPMLGFGAAGSTMGATMTL	VOARLLSGIVOOONMLLAIRAQOHL	566
Qy	529	QLTWGIKOQARVILVERTYLGDQQLG	IMGCSGKLICTAIVPMASNSNLSLDIRINNM	588
Db	567	QLTWGIKOQARVILVERTYLGDQQLG	IMGCSGKLICTTIVPMNTSMNSNLSLDIRINNM	622
Qy	589	TWMEWEEDINYSSELYTLLESQNOOEKNOB	ELLBDKMASLMMWEPITTWLWY	643
Db	627	TWMEWEEDINYSLLYTLLESQNOOEKNOB	ELLBDKMASLMMWEPITTWLWY	681

```

: INFORMATION FOR SEQ ID NO: 30:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 857 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: US-08-448-603A-30

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/134.075
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/448,603
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Halliday, Emily
REGISTRATION NUMBER: 38903
REFERENCE/DOCKET NUMBER: 14918-704
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-393-2000
TELEFAX: 415-393-2286
TELEX:
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 857 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-134-075-30

```

```

Query Match      86.6%; Score 2999; DB 3; Length 857;
Best Local Similarity 85.4%; Pred. No. 1.7e-242;
Matches 556; Conservative 41; Mismatches 44; Indels 10; Gaps 3;

```

```

QY 3 KLMVTVVYGVPMVKKATTTLLFCASDAKAYDREVNVAATHACVPTDPNPQEVLENTVEH 62
DB 32 KLMVTVVYGVPMVKKATTTLLFCASDAKAYDREVNVAATHACVPTDPNPQEVLENTVEH 91
QY 63 FNMKNKNVVEQMBEDIIISLMDQSLKPCVKLTPLCVTLNCKDV-NATNTTNDSEGMERGE 121
DB 92 FNMKNKNVVEQMBEDIIISLMDQSLKPCVKLTPLCVTLNCKDV-NATNTTNDSEGMERGE 151
QY 122 IKNSCFNITTSIRDEVQKAYALFYKLDVVPIDN-----NNTSYRLISCOTSVITTOACP 174
DB 152 IKNSCFNITTSIRDEVQKAYALFYKLDVVPIDN-----NNTSYRLISCOTSVITTOACP 211
QY 175 KISFEPPIHYCAPAGFALILKCNKDTFNGSKPCKNVSTVQCTHGIRPVVSTOLLNGSLA 234
DB 212 KISFEPPIHYCAPAGFALILKCNKDTFNGSKPCKNVSTVQCTHGIRPVVSTOLLNGSLA 271
QY 235 EEEVVISRDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGGRAFYTTGELIGDIR 294
DB 272 EEEVVISRDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGGRAFYTTGELIGDIR 331
QY 295 QAHCHISRAKNDTLKQIYIKURBOPENKTIIVFNHSSGGDPEIIVNHSFNGCEGFYCNST 354
DB 332 QAHCHISRAKNDTLKQIYIKURBOPENKTIIVFNHSSGGDPEIIVNHSFNGCEGFYCNST 391
QY 355 QLPNSTNNNN--TEGSSNTEGNTITLPCRIKQIIMMOEYKAMAPPRGQIRROSSNT 412
DB 392 QLPNSTNNNN--TEGSSNTEGNTITLPCRIKQIIMMOEYKAMAPPRGQIRROSSNT 451
QY 413 GLLLTRDGINENGTEIFRPGCGDMRDWRSEFYKYKVKVIEPLGVAFTKCRARYOREK 472
DB 452 GLLLTRDGINENGTEIFRPGCGDMRDWRSEFYKYKVKVIEPLGVAFTKCRARYOREK 511
QY 473 RAVGIGAVFLGFLGAAGSTMGAASMTLTYQARLLISGIYQOQNNILRAIEAOQRMQLTV 532
DB 512 RAVGIGAVFLGFLGAAGSTMGAASMTLTYQARLLISGIYQOQNNILRAIEAOQRMQLTV 571
QY 533 WGIKLOARVLAVERVYLDQOOLIGIWGSGKLIICCAAPVPMNWSMKSILDRIMNNMTWME 592
DB 572 WGIKLOARVLAVERVYLDQOOLIGIWGSGKLIICCAAPVPMNWSMKSILDRIMNNMTWME 631
QY 593 WEREIDNTSEIYTLIEESQNOQEKNEOBLLBLDKWASLMMNPDITNNLMWY 643

```

```

DB 632 WEREIDNTSEIYTLIEESQNOQEKNEOBLLBLDKWASLMMNPDITNNLMWY 682

```

```

RESULT 10
US-09-492-739-30
Sequence 30, Application US/09492739
Patent No. 6331404
GENERAL INFORMATION:
APPLICANT: Berman, Phillip W.
Nakanura, Gerald R.
TITLE OF INVENTION: HIV Envelope Polypeptides
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: McCutchen, Doyle, Brown & Enersen, LLP
STREET: 3 Embarcadero Center
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/492,739
FILING DATE: 27-Jan-2000
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/134,075
FILING DATE: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Halliday, Emily
REGISTRATION NUMBER: 38903
REFERENCE/DOCKET NUMBER: 14918-704
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-393-2000
TELEFAX: 415-393-2286
TELEX: <unknown>
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 857 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 30:
US-09-492-739-30

```

```

Query Match      86.6%; Score 2999; DB 4; Length 857;
Best Local Similarity 85.4%; Pred. No. 1.7e-242;
Matches 556; Conservative 41; Mismatches 44; Indels 10; Gaps 3;

```

```

QY 3 KLMVTVVYGVPMVKKATTTLLFCASDAKAYDREVNVAATHACVPTDPNPQEVLENTVEH 62
DB 32 KLMVTVVYGVPMVKKATTTLLFCASDAKAYDREVNVAATHACVPTDPNPQEVLENTVEH 91
QY 63 FNMKNKNVVEQMBEDIIISLMDQSLKPCVKLTPLCVTLNCKDV-NATNTTNDSEGMERGE 121
DB 92 FNMKNKNVVEQMBEDIIISLMDQSLKPCVKLTPLCVTLNCKDV-NATNTTNDSEGMERGE 151
QY 122 IKNSCFNITTSIRDEVQKAYALFYKLDVVPIDN-----NNTSYRLISCOTSVITTOACP 174
DB 152 IKNSCFNITTSIRDEVQKAYALFYKLDVVPIDN-----NNTSYRLISCOTSVITTOACP 211
QY 175 KISFEPPIHYCAPAGFALILKCNKDTFNGSKPCKNVSTVQCTHGIRPVVSTOLLNGSLA 234
DB 212 KISFEPPIHYCAPAGFALILKCNKDTFNGSKPCKNVSTVQCTHGIRPVVSTOLLNGSLA 271
QY 235 EEEVVISRDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGGRAFYTTGELIGDIR 294
DB 272 EEEVVISRDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGGRAFYTTGELIGDIR 331

```

```

Qy 295 QAHNCISRAKNDTLKQIVIKLREOFENKTIIVFNHSSGDEBEIVMHSNCEGEEFYCNST 354
Dy 332 QAHNCISRTDWNNTLQIVKLEQFGNKTIIPHNHSGGDEBEIVMHSNCEGEEFYCNST 391
Qy 355 QLFNSTMNN--TEGSNTEGNTITLPCRIOIIMMOEVBKAMAPPIRQIICSSNIT 412
Dy 392 QLFSTMTNTVSNSTSTENSTITLPCRIOIIMMOEVBKAMAPPIRQIICSSNIT 451
Qy 413 GLILTRDGINENGTEIFRPGGDMRDWRSEFYKYVKXIEPLGVAFTKCKRRVOREK 472
Dy 452 GLILTRDGSNNMNETRPGGDMRDWRSELYKYVKXIEPLGVAFTKCKRRVOREK 511
Qy 473 RAVSIGAVFLGFLGAAGSTMGASMTLTVQARLLLSGIVQOONNLRAIEAQOHLQITV 532
Dy 512 RAVGIGAVFLGFLGAAGSTMGASMTLTVQARLLLSGIVQOONNLRAIEAQOHLQITV 571
Qy 533 WGIQOLQARVLAVERYLGDQOLLGWGSGLICTTAVPMNASNSKSLDIRNNMTWME 592
Dy 572 WGIQOLQARVLAVERYLGDQOLLGWGSGLICTTAVPMNASNSKSLDIRNNMTWME 631
Qy 593 WEREIDNTSEIYTLIEESONQOEKNEBELLDKMAISLMMWPDITNMLWY 643
Dy 632 WEREIDNTSEIYTLIEESONQOEKNEBELLDKMAISLMMWPDITNMLWY 682

```

```

RESULT 11
US-08-127-499A-14
; Sequence 14, Application US/08127499A
; Patent No. 5510264
; GENERAL INFORMATION:
; APPLICANT: VAN ALSTYNE, Diane
; TITLE OF INVENTION: SHARMA, Lawrence Rajendra
; TITLE OF INVENTION: ANTIBODIES WHICH BIND MENINGITIS RELATED
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/127,499A
; FILING DATE: 28-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 51916/102/INBI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 861 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; US-08-127-499A-14

```

```

Query Match 86.4%; Score 2992.5; DB 1; Length 861;
Best Local Similarity 86.7%; Pred. No. 5.9e-242;
Matches 569; Conservative 32; Mismatches 40; Indels 15; Gaps 8;
Qy 2 EKLMTVTVYGVVWKEATITLFCASDAKAYTEVHANNWATHACVPTDPNPOEVLNVATE 61
Dy 32 EKLMTVTVYGVVWKEATITLFCASDAKAYTEVHANNWATHACVPTDPNPOEVLNVATE 91

```

```

Qy 62 HFNMMKNMTEOMQOEIIISLPDQSLKPCVLTPLCVTLNCKDY-NATNT-----TNDSG- 115
Dy 92 NFNMMKNMTEOMQOEIIISLPDQSLKPCVLTPLCVTLNCKDY-NATNT-----TNDSG- 151
Qy 116 -TMEGEIKKCSFNNTTSINDEVOKEYALFYKLDVPIIDNNNTSYRLISCDTSVITQACP 174
Dy 152 MMEKEIKKCSFNNTTSINDEVOKEYALFYKLDVPIIDNNNTSYRLISCDTSVITQACP 211
Qy 175 KISPEPIPIHYCAPAGAILKCKDKTFNGKPCKNSTVQCTHGIRPVYSTOLLNGSLA 234
Dy 212 KISPEPIPIHYCAPAGAILKCKDKTFNGKPCKNSTVQCTHGIRPVYSTOLLNGSLA 271
Qy 235 EEEVIRSDNFTNNAKTIIVQLKESVEINCTRNNTKRSIHT--GPRGAYTTGELIGD 292
Dy 272 EEEVIRSDNFTNNAKTIIVQLKESVEINCTRNNTKRSIHT--GPRGAYTTGELIGD 330
Qy 293 IROAHNCISRAKNDTLKQIVIKLREOF-ENKTIIVFNHSSGDEBEIVMHSNCEGEEFYC 351
Dy 331 IROAHNCISRAKNDTLKQIVIKLREOF-ENKTIIVFNHSSGDEBEIVMHSNCEGEEFYC 390
Qy 352 NSTOLFNSTMNN--TEGSNTEG-NTITLPCRIOIIMMOEVBKAMAPPIRQIIC 407
Dy 391 NSTOLFNSTMNN--TEGSNTEG-NTITLPCRIOIIMMOEVBKAMAPPIRQIIC 450
Qy 408 SSNITGLILTRDGINENGTEIFRPGGDMRDWRSEFYKYVKXIEPLGVAFTKCKRRV 467
Dy 451 SSNITGLILTRDGINENGTEIFRPGGDMRDWRSELYKYVKXIEPLGVAFTKCKRRV 510
Qy 468 VOREKRAVIGAVFLGFLGAAGSTMGASMTLTVQARLLLSGIVQOONNLRAIEAQOHL 527
Dy 511 VOREKRAVIGAVFLGFLGAAGSTMGASMTLTVQARLLLSGIVQOONNLRAIEAQOHL 570
Qy 528 LQITWGIQOLQARVLAVERYLGDQOLLGWGSGLICTTAVPMNASNSKSLDIRNNMTWME 587
Dy 571 LQITWGIQOLQARVLAVERYLGDQOLLGWGSGLICTTAVPMNASNSKSLDIRNNMTWME 630
Qy 588 MTEMEREIDNTSEIYTLIEESONQOEKNEBELLDKMAISLMMWPDITNMLWY 643
Dy 631 MTEMEREIDNTSEIYTLIEESONQOEKNEBELLDKMAISLMMWPDITNMLWY 686

```

```

RESULT 12
US-08-482-847-14
; Sequence 14, Application US/08482847
; Patent No. 5556757
; GENERAL INFORMATION:
; APPLICANT: VAN ALSTYNE, Diane
; TITLE OF INVENTION: SHARMA, Lawrence Rajendra
; TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPE SITES FOR
; TITLE OF INVENTION: BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,847
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/127,499
; FILING DATE: 28-SEP-1993
; ATTORNEY/AGENT INFORMATION:

```

NAME: BENT, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 51916/104/INBI
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 672-5300
 TELEFAX: (202) 672-5399
 TELEX: 904136
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 861 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: unknown
 US-08-482-847-14

Query Match 86.4%; Score 2992.5; DB 1; Length 861;
 Best Local Similarity 86.7%; Pred. No. 5.9e-242;
 Matches 569; Conservative 32; Mismatches 40; Indels 15; Gaps 8;

```

2 EKLWTVVYGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTE 61
32 EKLWTVVYGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTE 91
62 HNNMKNMVEQMOEDIIISLWDSLKPCVKLTPLCVTLNCKDV-NATNT---TNDSEG- 115
92 NFNMMKNDMVEQMHEDIISLWDSLKPCVKLTPLCVSLKCTDLGNATVNTNSNTSSSGE 151
116 -TMRGEIKNCSFNITTSIRBEVQKAYALFYKLDVVPIDNNNTSYRLISCTSVITQACP 174
152 MMEKEGELKNCSFNISTIRGVQKAYAFYKLDIIPIDNDTSTYRLISCTSVITQACP 211
175 KISFEPIPIHYCAPAGFAILKCNKDTFNGKPCPKVSTVOCTHGIRPVVSTOLLNGSLA 234
212 KVSFEPIPIHYCAPAGFAILKCNKDTFNGKPCPKVSTVOCTHGIRPVVSTOLLNGSLA 271
235 EEEVIRSDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIH--GPRAFYTTGEIIGD 292
272 EEEVIRSANFTNNAKTIIVQLNQSVEINCTRPNNNTRKSIIRIORGPRAFTVIGK-IGN 330
293 IROAHCHNISRAKANDTLKQIVYKLRQF-EKNTIVFNHSSGGDEPIVHNSNGCEFPYC 351
331 MROAHCHNISRAKANDTLKQIVYKLRQF-EKNTIVFNHSSGGDEPIVHNSNGCEFPYC 390
352 NSTOLFSTWNN---NTEGSNNTG-NITTLPCRIRKQIINMVEGKAMVAPPIRGQIRG 407
391 NSTOLFSTWNNSTWNTSGSNTBSDTITLPCRIRKQIINMVEGKAMVAPPIRGQIRG 450
408 SSNITGLLTRDGGINENGTETFRPGGDMRDMNRSEFYKVKVYKIEPLGVAFTKCKRRV 467
451 SSNITGLLTRDGGINENGTETFRPGGDMRDMNRSEFYKVKVYKIEPLGVAFTKCKRRV 510
468 VOREBRAVIGAVPLGFIAGAAGSTMGASMTLTVQARLLSGIVQOONNLRAIEAQORM 527
511 VOREBRAVIGAVPLGFIAGAAGSTMGASMTLTVQARLLSGIVQOONNLRAIEAQOHL 570
528 LQLTWGKIQOLQARLAVERYLQDQQLGIMGCSKLLCTTAVPNNASMSKSLDRIRNN 587
571 LQLTWGKIQOLQARLAVERYLQDQQLGIMGCSKLLCTTAVPNNASMSKSLDRIRNN 630
588 MTMMEWEREDNYSYSEIYTLIEESQOQEKNOQELLEDKVASLWMDITNMLWY 643
631 MTMMEWEREDNYSYSEIYTLIEESQOQEKNOQELLEDKVASLWMDITNMLWY 686

```

RESULT 13
 US-07-956-483-10
 Sequence 10, Application US/07956483
 Patent No. 6261799
 GENERAL INFORMATION:
 APPLICANT: KIENY, Marie-Paule
 TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED
 TITLE OF INVENTION: gp160 VARIANT
 NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Burns, Doane, Swecker & Mathis
 STREET: P. O. Box 1404
 CITY: Alexandria
 STATE: Virginia
 COUNTRY: United States
 ZIP: 22313-1404
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/956,483
 FILING DATE: 31-DEC-1992
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO 92/19742
 FILING DATE: 12-NOV-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: FR 91 05392
 FILING DATE: 02-MAY-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Crane-Feury, Sharon E
 REGISTRATION NUMBER: 36,113
 REFERENCE/DOCKET NUMBER: 01753-005
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 836-6620
 TELEFAX: (703) 836-2021
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 861 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 US-07-956-483-10

Query Match 86.4%; Score 2992.5; DB 3; Length 861;
 Best Local Similarity 86.7%; Pred. No. 5.9e-242;
 Matches 569; Conservative 32; Mismatches 40; Indels 15; Gaps 8;

```

2 EKLWTVVYGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTE 61
32 EKLWTVVYGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTE 91
62 HNNMKNMVEQMOEDIIISLWDSLKPCVKLTPLCVTLNCKDV-NATNT---TNDSEG- 115
92 NFNMMKNDMVEQMHEDIISLWDSLKPCVKLTPLCVSLKCTDLGNATVNTNSNTSSSGE 151
116 -TMRGEIKNCSFNITTSIRBEVQKAYALFYKLDVVPIDNNNTSYRLISCTSVITQACP 174
152 MMEKEGELKNCSFNISTIRGVQKAYAFYKLDIIPIDNDTSTYRLISCTSVITQACP 211
175 KISFEPIPIHYCAPAGFAILKCNKDTFNGKPCPKVSTVOCTHGIRPVVSTOLLNGSLA 234
212 KVSFEPIPIHYCAPAGFAILKCNKDTFNGKPCPKVSTVOCTHGIRPVVSTOLLNGSLA 271
235 EEEVIRSDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIH--GPRAFYTTGEIIGD 292
272 EEEVIRSANFTNNAKTIIVQLNQSVEINCTRPNNNTRKSIIRIORGPRAFTVIGK-IGN 330
293 IROAHCHNISRAKANDTLKQIVYKLRQF-EKNTIVFNHSSGGDEPIVHNSNGCEFPYC 351
331 MROAHCHNISRAKANDTLKQIVYKLRQF-EKNTIVFNHSSGGDEPIVHNSNGCEFPYC 390
352 NSTOLFSTWNN---NTEGSNNTG-NITTLPCRIRKQIINMVEGKAMVAPPIRGQIRG 407
391 NSTOLFSTWNNSTWNTSGSNTBSDTITLPCRIRKQIINMVEGKAMVAPPIRGQIRG 450
408 SSNITGLLTRDGGINENGTETFRPGGDMRDMNRSEFYKVKVYKIEPLGVAFTKCKRRV 467
451 SSNITGLLTRDGGINENGTETFRPGGDMRDMNRSEFYKVKVYKIEPLGVAFTKCKRRV 510

```


Qy 468 VOREKRAVGIGAVFLGFGAAGSTMGASMTLTVQARLLLSGIYQOONNLRAIEAQQRM 527
 Db 511 VOREKRAVGIGAVFLGFGAAGSTMGASMTLTVQARLLLSGIYQOONNLRAIEAQQHL 570
 Qy 528 LQITWVGIKOLQARVLAVERYLDQOOLLGIVGSGSKLICCTAVPWNASWSNKSIDRIWNN 587
 Db 571 LQITWVGIKOLQARVLAVERYLDQOOLLGIVGSGSKLICCTAVPWNASWSNKSIEQIMNN 630
 Qy 588 MTWMEWREIDNTSEIYTLIEESONQOEKNEBELLEDKASLMMNFDITNNMLY 643
 Db 631 MTWMEWREIDNTSEIYTLIEESONQOEKNEBELLEDKASLMMNFDITNNMLY 686

RESULT 14

US-08-472-240A-1
 ; Sequence 1, Application US/08472240A
 ; Patent No. 6284248
 ; GENERAL INFORMATION:
 APPLICANT: KIENY, Marie-Paule
 TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED
 NUMBER OF SEQUENCES: 29
 ; CORRESPONDENCE ADDRESS:
 ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
 STREET: P.O. Box 1404
 CITY: Alexandria
 STATE: Virginia
 COUNTRY: United States
 ZIP: 22313-1404
 ; COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/472.240A
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/956,483
 FILING DATE: 31-DEC-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Teskin, Robin L.
 REGISTRATION NUMBER: 35,030
 REFERENCE/DOCKET NUMBER: 017753-055
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 836-6620
 TELEFAX: (703) 836-6620
 ; INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 661 amino acids
 TYPE: amino acid
 STRANDEDNESS: not relevant
 TOPOLOGY: not relevant
 MOLECULE TYPE: peptide
 FEATURE:
 NAME/KEY: Protein
 LOCATION: 1..854
 ; US-08-472-240A-1

Query March 86.4%; Score 2992.5; DB 3; Length 661;
 Best Local Similarity 86.7%; Pred. No. 5.9e-242;
 Matches 569; Conservative .32; Mismatches 40; Indels 15; Gaps 8;
 Qy 2 EKLWTVYVGVVWKEATTTTLCFASDAKAYDTEVHNVWATHACVPTDPNPOEVLEVENTE 61
 Db 32 EKLWTVYVGVVWKEATTTTLCFASDAKAYDTEVHNVWATHACVPTDPNPOEVLVANTE 91
 Qy 62 HFNWKNMNVQMOEDITSLWDQSLKPCVKLTPLCVTLNCKDV-NATNT---TNDSFG- 115
 Db 92 NFNMKQDMVQMOEDITSLWDQSLKPCVKLTPLCVSLKCTDLDGNATNTSSNTSSSGE 151

Qy 116 -TWERGEIKNCSFNITTSIRDEVOKEAYFLKLVVPIIDNNNTSYRLISCDTSVITQACP 174
 Db 152 MMKEGEBIKNCSFNITTSIRGKVQKEAFYKLDIIPIDNDTTSYRLISCDTSVITQACP 211
 Qy 175 KISFEPIPIHYCAPAGFAILKCDKTPNGKPCANSTVQCTGIRPVVSTOLLNGSLA 234
 Db 212 KVSFEPIPIHYCAPAGFAILKCNKTFNGTGPCTNVSTVQCTGIRPVVSTOLLNGSLA 271
 Qy 235 EEEVVRISDNPFTNNAKTIIVQLKESVEINCTRENNNTSKSIH--GPRGFYTTGEIIGD 292
 Db 272 EEEVVRISANFTDPAKTIIVQLNQSVEINCTRENNNTSKSIRLORGFRAFTIIGK-IGN 330
 Qy 293 IRQAHCNISRAKNNDTLQIVILKEQF-BNKTIIVNHSQSGDEIYVHSGFNGEPEFYC 351
 Db 331 MRQAHCNISRAKNNDTLQIVILKEQF-BNKTIIVNHSQSGDEIYVHSGFNGEPEFYC 390
 Qy 352 NSTOLFNSFTWNN--NTGSGNTEG-NTTLPCRIQIIMMOEYKAMAPPIRGQIRC 407
 Db 391 NSTOLFNSFTWNN--NTGSGNTEG-NTTLPCRIQIIMMOEYKAMAPPIRGQIRC 450
 Qy 408 SSNTGLLTRDGINENGTEIFRPGGDMRDMRSFYYKYVYKIEPLGVAPFKGRV 467
 Db 451 SSNTGLLTRDGINENGTEIFRPGGDMRDMRSFYYKYVYKIEPLGVAPFKGRV 510
 Qy 468 VOREKRAVGIGAVFLGFGAAGSTMGASMTLTVQARLLLSGIYQOONNLRAIEAQQRM 527
 Db 511 VOREKRAVGIGAVFLGFGAAGSTMGASMTLTVQARLLLSGIYQOONNLRAIEAQQHL 570
 Qy 528 LQITWVGIKOLQARVLAVERYLDQOOLLGIVGSGSKLICCTAVPWNASWSNKSIDRIWNN 587
 Db 571 LQITWVGIKOLQARVLAVERYLDQOOLLGIVGSGSKLICCTAVPWNASWSNKSIEQIMNN 630
 Qy 588 MTWMEWREIDNTSEIYTLIEESONQOEKNEBELLEDKASLMMNFDITNNMLY 643
 Db 631 MTWMEWREIDNTSEIYTLIEESONQOEKNEBELLEDKASLMMNFDITNNMLY 686

RESULT 15

US-08-472-240A-7
 ; Sequence 7, Application US/08472240A
 ; Patent No. 6284248
 ; GENERAL INFORMATION:
 APPLICANT: KIENY, Marie-Paule
 TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED
 NUMBER OF SEQUENCES: 29
 ; CORRESPONDENCE ADDRESS:
 ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
 STREET: P.O. Box 1404
 CITY: Alexandria
 STATE: Virginia
 COUNTRY: United States
 ZIP: 22313-1404
 ; COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/472.240A
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/956,483
 FILING DATE: 31-DEC-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Teskin, Robin L.
 REGISTRATION NUMBER: 35,030
 REFERENCE/DOCKET NUMBER: 017753-055
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 836-6620
 TELEFAX: (703) 836-6620
 ; INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:
 LENGTH: 861 amino acids
 TYPE: amino acid
 STRANDEDNESS: not relevant
 TOPOLOGY: not relevant
 MOLECULE TYPE: peptide
 FEATURE:
 NAME/KEY: Protein
 LOCATION: 1..831
 US-08-472-240A-7

Query Match 86.4%; Score 2992.5; DB 3; Length 861;
 Best Local Similarity 86.7%; Pred. No. 5.9e-242;
 Matches 569; Conservative 32; Mismatches 40; Indels 15; Gaps 8;

```

QY 2 EKLMTVTVYGVVPMWEATTTTLFCASDADAYDTEVHNWATHACVPTDPNPOEVLENTE 61
DB 32 EKLMTVTVYGVVPMWEATTTTLFCASDADAYDTEVHNWATHACVPTDPNPOEVLENTE 91
62 HFNMMKNNMVEQMORDIISLMDQSLKPCVKLTPLCVTLNCKDV-NATNT---TNDSEG- 115
92 NFNMMKNNMVEQMORDIISLMDQSLKPCVKLTPLCVSLKCTDLGNATNTNSNTSSGE 151
116 -TMRGEIKNCSEFNITTSIRDEVEKEVALFYKLDVVPIDNNNTSYRLISCDTSVITQACP 174
152 MMEKEGEEKNCSEFNITTSIRGEVKEVAFYKLDIIPIDNPTSYTLTSCNTSVITQACP 211
175 KISFEPIPIHVCAPGPAFLKNDKTFNGKGPCKAVSTVOCTHGIRPVVSTOLLNGSLA 234
212 KVSFSPPIPIHVCAPGPAFLKCNKTFNGTGTCTVSTVOCTHGIRPVVSTOLLNGSLA 271
235 EEEVYIRSDNFTNNAKTIIIVOLKESVEINCTRPNNNTRKSIH--GPGRAFTTGEITGD 292
272 EEEVYIRSANFTDNKTIIVOLNGSVEINCTRPNNNTRKSIH--GPGRAFTTGEITGD 330
293 IROAHCNISRAKMDTLKQIVIKLEQF-ENKTIIVNHSGGDPEIIVMHSFNCGESEFFYC 351
331 MROAHCNISRAKMDTLKQIVIKLEQF--GNNKTIIFKQSSGGDPEIIVHSGGESEFFYC 390
352 NSTOLFNGSTNN--NTGSSNTEG-NTITLPCRIKQIINMQVYKAMYPAPIRGQIRC 407
391 NSTOLFNGSTNNTEGSDTITLPCRIKQIINMQVYKAMYPAPIRGQIRC 450
408 SSNTGLLTLTRDGINENGTEIFRPGGDMRDNRSEFYKYVVKIEPIGVAPTCKKRV 467
451 SSNTGLLTLTRDGINNGNSEIFRPGGDMRDNRSELYKYVVKIEPIGVAPTAKRV 510
468 VOREIRAVGIGAVFLGFLGAAGSTWGAASMTLTVQARLLSGIVQOQNNLLRAIEAQRM 527
511 VOREIRAVGIGALFLGFLGAAGSTWGAASMTLTVQARLLSGIVQOQNNLLRAIEAQHL 570
528 LQLTWGIKQLOARVLAVERYIGDOOLLGIWCGSKLICCTAVPWNASWSNKSLEDRINWN 587
571 LQLTWGIKQLOARVLAVERYIKDOOLLGIWCGSKLICCTAVPWNASWSNKSLEQIWN 630
588 MTWMEBREINDYTSEIYTLIEESONQOEKNEBELLEDKNASLNNWFDITWMLY 643
631 MTWMEBREINDYTSEIYTLIEESONQOEKNEBELLEDKNASLNNWFDITWMLY 686

```

Search completed: December 12, 2003, 12:34:26
 Job time : 16.6215 secs

THIS PAGE BLANK (US 710)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 12:32:44 ; Search time 26.8062 Seconds

(without alignments)
4461.192 Million cell updates/sec

Title: US-10-032-162-13

Perfect score: 3462
Sequence: 1 VEKLMVTYVYGVVPWKEATT.....ELDKMASIMWFDITNMLWY 643

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

1 number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCF_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/CTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10C_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3462	100.0	643	US-10-032-162-13	Sequence 13, Appl
2	3295	95.2	625	US-10-032-162-17	Sequence 17, Appl
3	3056	88.3	579	US-10-032-162-15	Sequence 15, Appl
4	3041	87.8	842	US-10-190-435-2	Sequence 2, Appl
5	3041	87.8	842	US-10-241-009-2	Sequence 2, Appl
6	3041	87.8	842	US-10-190-434B-2	Sequence 2, Appl
7	3041	87.8	842	US-10-190-305A-2	Sequence 2, Appl
8	3041	87.8	847	US-09-476-242-2	Sequence 2, Appl
9	2992.5	86.4	861	US-10-026-741-103	Sequence 103, App
10	2964	85.6	856	US-09-476-242-1	Sequence 1, Appl
11	2953	85.3	856	US-10-196-515-11	Sequence 11, Appl
12	2910.5	84.1	726	US-10-196-515-3	Sequence 3, Appl
13	2897.5	83.7	759	US-10-196-515-12	Sequence 12, Appl
14	2895.5	83.6	868	US-09-938-406-1	Sequence 1, Appl
15	2862	82.7	619	US-09-891-609-4	Sequence 4, Appl

16	2862	82.7	646	US-09-891-609-2	Sequence 2, Appl
17	2736.5	79.0	860	US-10-190-435-6	Sequence 6, Appl
18	2736.5	79.0	860	US-10-241-009-6	Sequence 6, Appl
19	2736.5	79.0	860	US-10-190-434B-6	Sequence 6, Appl
20	2736.5	79.0	860	US-10-190-305A-6	Sequence 6, Appl
21	2696	77.9	867	US-10-190-435-3	Sequence 3, Appl
22	2696	77.9	867	US-10-190-435-126	Sequence 126, App
23	2696	77.9	867	US-10-241-009-3	Sequence 3, Appl
24	2696	77.9	867	US-10-190-434B-3	Sequence 3, Appl
25	2696	77.9	867	US-10-190-305A-3	Sequence 3, Appl
26	2694	77.8	858	US-10-190-435-150	Sequence 150, App
27	2687	77.6	869	US-10-190-435-4	Sequence 4, Appl
28	2687	77.6	869	US-10-190-434B-4	Sequence 4, Appl
29	2687	77.6	869	US-10-241-009-4	Sequence 4, Appl
30	2687	77.6	869	US-10-190-305A-4	Sequence 4, Appl
31	2684.5	77.5	803	US-10-190-435-135	Sequence 135, App
32	2683.5	77.5	803	US-10-190-435-134	Sequence 134, App
33	2683.5	77.5	870	US-10-190-435-127	Sequence 127, App
34	2682	77.5	861	US-10-190-435-139	Sequence 139, App
35	2670	77.1	870	US-10-190-435-147	Sequence 147, App
36	2669	77.1	862	US-10-190-435-141	Sequence 141, App
37	2668.5	77.1	865	US-10-190-435-140	Sequence 140, App
38	2668.5	77.0	857	US-10-190-435-138	Sequence 138, App
39	2666	77.0	849	US-10-190-435-148	Sequence 148, App
40	2662.5	76.9	855	US-10-190-435-144	Sequence 144, App
41	2660.5	76.8	845	US-10-190-435-129	Sequence 129, App
42	2660.5	76.8	845	US-10-190-435-130	Sequence 130, App
43	2648	76.5	855	US-10-369-294-11	Sequence 11, Appl
44	2648	76.5	855	US-10-369-294-12	Sequence 12, Appl
45	2644	76.4	844	US-09-991-258-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1:
US-10-032-162-13
; Sequence 13, Application US/10032162
; Publication No. US20030052839A1
GENERAL INFORMATION:
; APPLICANT: BINLEY, JAMES M
; APPLICANT: SCHUELEKE, NORBERT
; APPLICANT: OLSON, WILLIAM C
; APPLICANT: PAUL, MADON J
; APPLICANT: JOHN, MOORE P
TITLE OR INVENTION: STABILIZED VIRAL ENVELOPE PROTEINS AND USES THEREOF
; FILE REFERENCE: 2048/59331az
; CURRENT APPLICATION NUMBER: US/10/032,162
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/602,864
; PRIOR FILING DATE: 2000-06-23
NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
SEQ ID NO 13
; LENGTH: 643
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-10-032-162-13

Query Match 100.0%; Score 3462; DB 15; Length 643;
Best Local Similarity 100.0%; Pred. No. 1.5e-308;
Matches 643; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEKLMVTYVYGVVPWKEATTTLFCASDAKAYDTEVHNHVAATHACVPTDPNQEVLLENT 60
DB 1 VEKLMVTYVYGVVPWKEATTTLFCASDAKAYDTEVHNHVAATHACVPTDPNQEVLLENT 60
QY EHFMMKNNNVEQOEEDIIISLMDOSLKPQVGLPTCLNCKDYNATNTTDSGTERG 120
DB EHFMMKNNNVEQOEEDIIISLMDOSLKPQVGLPTCLNCKDYNATNTTDSGTERG 120
QY EIKKSNITTSIRDEVOKEAYALFYKLDVVPIDNNNTSYRLISCDTSVITQACPKISFEP 180

```

Db      121 EIKNCSEFNITTSIRDEVOKEYALFYKLDVVPIDNNNTSYRLISCDTSVITQACPXISFEP 180
Qy      181 IPIHYCAPAGFALIKNDKTFNGKGPCKNVSTVOCTHGIRPVYSTOLLNGSLAEERVI 240
Db      181 IPIHYCAPAGFALIKNDKTFNGKGPCKNVSTVOCTHGIRPVYSTOLLNGSLAEERVI 240
Qy      241 RSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSHIGPRAFYTTGEIIGDIRQAHANI 300
Db      241 RSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSHIGPRAFYTTGEIIGDIRQAHANI 300
Qy      301 SRKAMDITLKOIYIKLREOPENKTIIVFNHSSGDPPEIWMHSFNCGEFFFCNSTQULFNST 360
Db      301 SRKAMDITLKOIYIKLREOPENKTIIVFNHSSGDPPEIWMHSFNCGEFFFCNSTQULFNST 360
Qy      361 MNNTGSSNTEGNTITLPCRIKOIIMMOEVGKAMAPPIRGQIRCSSNITGLLTRDG 420
Db      361 MNNTGSSNTEGNTITLPCRIKOIIMMOEVGKAMAPPIRGQIRCSSNITGLLTRDG 420
Qy      421 GINENGTEIFRPGGDMRWSEFYKYKVKIEPLGVAPTKCRVVOREKRAVGIGAV 480
Db      421 GINENGTEIFRPGGDMRWSEFYKYKVKIEPLGVAPTKCRVVOREKRAVGIGAV 480
Qy      481 FLGFLGAAGSTMAASNTLTVQARLLISGIVQOONNLLRAIEAQORMLQLTWNGIKOLOA 540
Db      481 FLGFLGAAGSTMAASNTLTVQARLLISGIVQOONNLLRAIEAQORMLQLTWNGIKOLOA 540
Qy      541 RVLAVERYLDGQDLGIGWGSGLICTTAVPMAASNSKSLDRIMNMNTWMEEREIDNY 600
Db      541 RVLAVERYLDGQDLGIGWGSGLICTTAVPMAASNSKSLDRIMNMNTWMEEREIDNY 600
Qy      601 TSEIYTLIESONQOEKNEQELLEDKMASLMMNFDITNMLWY 643
Db      601 TSEIYTLIESONQOEKNEQELLEDKMASLMMNFDITNMLWY 643

```

```

RESULT 2
US-10-032-162-17
; Sequence 17, Application US/10032162
; Publication No. US20030052839A1
; GENERAL INFORMATION:
; APPLICANT: BINEY, JAMES M
; APPLICANT: SCHUELE, NORBERT
; APPLICANT: OLSON, WILLIAM C
; APPLICANT: PAUL, MADDON J
; APPLICANT: JOHN, MOORE P
; TITLE OR INVENTION: STABILIZED VIRAL ENVELOPE PROTEINS AND USES THEREOF
; FILE REFERENCE: 2048/59331a2
; CURRENT APPLICATION NUMBER: US/10/032,162
; PRIOR FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/602,864
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 625
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (151)..(151)
; OTHER INFORMATION: X=UNKNOWN AMINO ACID
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (205)..(205)
; OTHER INFORMATION: X=UNKNOWN AMINO ACID
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (213)..(213)
; OTHER INFORMATION: X=UNKNOWN AMINO ACID
US-10-032-162-17

```

Query Match 95.2%; Score 3295; DB 15; Length 625;
Best Local Similarity 96.0%; Pred. No. 3.le-293;

```

Matches 617; Conservative 0; Mismatches 8; Indels 18; Gaps 1;
Qy      1 VEKLMATVYGVVWMEATATTLFCASDAXAYDEVHNVNATHACVPTDPPOEVLENT 60
Db      1 VEKLMATVYGVVWMEATATTLFCASDAXAYDEVHNVNATHACVPTDPPOEVLENT 60
Qy      61 EHNMMKNMNVMEQMDIISLMDQSLKPCVKLTPLCVTLNCKDVNATNTTNDSEGTMERG 120
Db      61 EHNMMKNMNVMEQMDIISLMDQSLKPCVKLTPLCVTLNCKDVNATNTTNDSEGTMERG 120
Qy      121 EIKNCSEFNITTSIRDEVOKEYALFYKLDVVPIDNNNTSYRLISCDTSVITQACPXISFEP 180
Db      121 EIKNCSEFNITTSIRDEVOKEYALFYKLDVVPIDNNNTSYRLISCDTSVITQACPXISFEP 180
Qy      181 IPIHYCAPAGFALIKNDKTFNGKGPCKNVSTVOCTHGIRPVYSTOLLNGSLAEERVI 240
Db      181 IPIHYCAPAGFALIKNDKTFNGKGPCKNVSTVOCTHGIRPVYSTOLLNGSLAEERVI 240
Qy      241 RSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSHIGPRAFYTTGEIIGDIRQAHANI 300
Db      241 RSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSHIGPRAFYTTGEIIGDIRQAHANI 300
Qy      301 SRKAMDITLKOIYIKLREOPENKTIIVFNHSSGDPPEIWMHSFNCGEFFFCNSTQULFNST 360
Db      301 SRKAMDITLKOIYIKLREOPENKTIIVFNHSSGDPPEIWMHSFNCGEFFFCNSTQULFNST 360
Qy      361 MNNTGSSNTEGNTITLPCRIKOIIMMOEVGKAMAPPIRGQIRCSSNITGLLTRDG 420
Db      361 MNNTGSSNTEGNTITLPCRIKOIIMMOEVGKAMAPPIRGQIRCSSNITGLLTRDG 420
Qy      421 GINENGTEIFRPGGDMRWSEFYKYKVKIEPLGVAPTKCRVVOREKRAVGIGAV 480
Db      421 GINENGTEIFRPGGDMRWSEFYKYKVKIEPLGVAPTKCRVVOREKRAVGIGAV 480
Qy      481 FLGFLGAAGSTMAASNTLTVQARLLISGIVQOONNLLRAIEAQORMLQLTWNGIKOLOA 540
Db      481 FLGFLGAAGSTMAASNTLTVQARLLISGIVQOONNLLRAIEAQORMLQLTWNGIKOLOA 540
Qy      541 RVLAVERYLDGQDLGIGWGSGLICTTAVPMAASNSKSLDRIMNMNTWMEEREIDNY 600
Db      541 RVLAVERYLDGQDLGIGWGSGLICTTAVPMAASNSKSLDRIMNMNTWMEEREIDNY 600
Qy      601 TSEIYTLIESONQOEKNEQELLEDKMASLMMNFDITNMLWY 643
Db      583 TSEIYTLIESONQOEKNEQELLEDKMASLMMNFDITNMLWY 625

```

```

RESULT 3
US-10-032-162-15
; Sequence 15, Application US/10032162
; Publication No. US20030052839A1
; GENERAL INFORMATION:
; APPLICANT: BINEY, JAMES M
; APPLICANT: SCHUELE, NORBERT
; APPLICANT: OLSON, WILLIAM C
; APPLICANT: PAUL, MADDON J
; APPLICANT: JOHN, MOORE P
; TITLE OR INVENTION: STABILIZED VIRAL ENVELOPE PROTEINS AND USES THEREOF
; FILE REFERENCE: 2048/59331a2
; CURRENT APPLICATION NUMBER: US/10/032,162
; PRIOR FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/602,864
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-10-032-162-15

```

Query Match 88.3%; Score 3056; DB 15; Length 579;
Best Local Similarity 89.3%; Pred. No. 2.3e-271;

Matches 574; Conservative 0; Mismatches 5; Indels 64; Gaps 1;

```

QY 1 VEKLMVTVYGVPPWKEATTTTLFCASDAKAYDEVHNMWATHACVPTDPNQEVLLENT 60
DB 1 VEKLMVTVYGVPPWKEATTTTLFCASDAKAYDEVHNMWATHACVPTDPNQEVLLENT 60
QY 61 EHPNMKNMWEQOEEDIIISLMDOSLKPVCVLTPLCTVLNCKQVATNTJNDSBGTMERG 120
DB 61 EHPNMKNMWEQOEEDIIISLMDOSLKPVCVLTPLCTVLNCKQVATNTJNDSBGTMERG 120
QY 121 EIKNCSFNITTSIRDEVOKEYALFYKLDVVPIDNNNTSYRLISCDTSVITQACPKISFE 180
DB 99 -----GCDTSVITQACPKISFE 116
QY 181 IPIHYCAPAGFALLKNDKTFNGKPCKNVSTVOCTHGIRPVVSTOLLNGSLAEDEV 240
DB 117 IPIHYCAPAGFALLKNDKTFNGKPCKNVSTVOCTHGIRPVVSTOLLNGSLAEDEV 176
QY 241 RSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGRAFYTTGELIGDROAHCN 300
DB 177 RSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGRAFYTTGELIGDROAHCN 236
QY 301 SRAKNNDTLKOIVIKLREOFENKTIIVFNHSSGDEPEIVMHSFGCEFFYCNSDTOLFNS 360
DB 237 SRAKNNDTLKOIVIKLREOFENKTIIVFNHSSGDEPEIVMHSFGCEFFYCNSDTOLFNS 296
QY 361 WNNNTEGNNTEGNTITLPCRKOIINMOEYKAMAPPIRGQIRCSSNITGLLTRD 420
DB 297 WNNNTEGNNTEGNTITLPCRKOIINMOEYKAMAPPIRGQIRCSSNITGLLTRD 356
QY 421 GINENGTEIFRGGGDMRDMRSEFYKYVVIIEPLGAPATCKRVRVOREKRAVIGAV 480
DB 357 GINENGTEIFRGGGDMRDMRSEFYKYVVIIEPLGAPATCKRVRVOREKRAVIGAV 416
QY 481 FLFLGAGASTWGASMTLTVQARLLSGIVQOQNNLLRAIEAQRMQLQTVWGIKOLA 540
DB 417 FLFLGAGASTWGASMTLTVQARLLSGIVQOQNNLLRAIEAQRMQLQTVWGIKOLA 476
QY 541 RVLAVERYLGDQOLGIGWCSGKLICTTAVPNNASMSKSLDRINNNMTWMEWEREID 600
DB 477 RVLAVERYLGDQOLGIGWCSGKLICTTAVPNNASMSKSLDRINNNMTWMEWEREID 536
QY 601 TSEIYTLIESONQOEKNEOELELDKMASLNNMFDITNMLY 643
DB 537 TSEIYTLIESONQOEKNEOELELDKMASLNNMFDITNMLY 579

```

RESULT 4

US-190-435-2

Sequence 2, Application US/10190435

Publication No. US20030143248A1

GENERAL INFORMATION:

APPLICANT: ZUR MEGEDE, Jan

APPLICANT: BARNETT, Susan W.

APPLICANT: LIAN, Ying

APPLICANT: ENGELBRECHT, Susan

APPLICANT: VAN RENSBURG, Estrelita J.

TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C

FILE REFERENCE: PPI18133.003 / 2302-18133

CURRENT APPLICATION NUMBER: US/10/190,435

NUMBER OF SEQ ID NOS: 319

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2

LENGTH: 842

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: SP162

US-10-190-435-2

Query Match 87.8%; Score 3041; DB 12; Length 842;

Best Local Similarity 88.7%; Pred. No. 9,66-270; Matches 572; Conservative 29; Mismatches 40; Indels 4; Gaps 4;

```

QY 1 VEKLMVTVYGVPPWKEATTTTLFCASDAKAYDEVHNMWATHACVPTDPNQEVLLENT 60
DB 1 VEKLMVTVYGVPPWKEATTTTLFCASDAKAYDEVHNMWATHACVPTDPNQEVLLENT 60
QY 25 VEKLMVTVYGVPPWKEATTTTLFCASDAKAYDEVHNMWATHACVPTDPNQEVLLENT 84
DB 25 VEKLMVTVYGVPPWKEATTTTLFCASDAKAYDEVHNMWATHACVPTDPNQEVLLENT 84
QY 61 EHPNMKNMWEQOEEDIIISLMDOSLKPVCVLTPLCTVLNCKQVATNTJNDSBGTMER 119
DB 61 EHPNMKNMWEQOEEDIIISLMDOSLKPVCVLTPLCTVLNCKQVATNTJNDSBGTMER 119
QY 85 EHPNMKNMWEQOEEDIIISLMDOSLKPVCVLTPLCTVLNCKQVATNTJNDSBGTMER 144
DB 85 EHPNMKNMWEQOEEDIIISLMDOSLKPVCVLTPLCTVLNCKQVATNTJNDSBGTMER 144
QY 120 GEIKNCSFNITTSIRDEVOKEYALFYKLDVVPIDNNNTSYRLISCDTSVITQACPKISFE 179
DB 145 GEIKNCSFNITTSIRDEVOKEYALFYKLDVVPIDNNNTSYRLISCDTSVITQACPKISFE 204
QY 180 PIPHYCAPAGFALLKNDKTFNGKPCKNVSTVOCTHGIRPVVSTOLLNGSLAEDEV 239
DB 205 PIPHYCAPAGFALLKNDKTFNGKPCKNVSTVOCTHGIRPVVSTOLLNGSLAEDEV 264
QY 240 IRSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGRAFYTTGELIGDROAHCN 299
DB 265 IRSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGRAFYTTGELIGDROAHCN 324
QY 300 ISRAKNNDTLKOIVIKLREOFENKTIIVFNHSSGDEPEIVMHSFGCEFFYCNSDTOLFNS 359
DB 325 ISRAKNNDTLKOIVIKLREOFENKTIIVFNHSSGDEPEIVMHSFGCEFFYCNSDTOLFNS 384
QY 360 TNNNTEGNNTEGNTITLPCRKOIINMOEYKAMAPPIRGQIRCSSNITGLLTRD 419
DB 385 TNNNTEGNNTEGNTITLPCRKOIINMOEYKAMAPPIRGQIRCSSNITGLLTRD 442
QY 420 GGIN-ENGTEIFRGGGDMRDMRSEFYKYVVIIEPLGAPATCKRVRVOREKRAVIGAV 478
DB 443 GGIN-ENGTEIFRGGGDMRDMRSEFYKYVVIIEPLGAPATCKRVRVOREKRAVIGAV 502
QY 479 AVFLGAGASTWGASMTLTVQARLLSGIVQOQNNLLRAIEAQRMQLQTVWGIKOLA 538
DB 503 AVFLGAGASTWGASMTLTVQARLLSGIVQOQNNLLRAIEAQRMQLQTVWGIKOLA 562
QY 539 QARVLAVERYLGDQOLGIGWCSGKLICTTAVPNNASMSKSLDRINNNMTWMEWEREID 598
DB 563 QARVLAVERYLGDQOLGIGWCSGKLICTTAVPNNASMSKSLDRINNNMTWMEWEREID 622
QY 599 NYTSEIYTLIESONQOEKNEOELELDKMASLNNMFDITNMLY 643
DB 623 NYTSEIYTLIESONQOEKNEOELELDKMASLNNMFDITNMLY 667

```

RESULT 5

US-10-241-009-2

Sequence 2, Application US/10241009

Publication No. US20030170614A1

GENERAL INFORMATION:

APPLICANT: ZUR MEGEDE, Jan

APPLICANT: BARNETT, Susan

APPLICANT: LIAN, Ying

APPLICANT: VAN RENSBURG, Estrelita J.

TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B

FILE REFERENCE: 2300-1621.21

CURRENT APPLICATION NUMBER: US/10/241,009

NUMBER OF SEQ ID NOS: 68

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2

LENGTH: 842

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

US-10-241-009-2

Query Match 87.8%; Score 3041; DB 12; Length 842;

Best Local Similarity 88.7%; Pred. No. 9,66-270;

Matches 572; Conservative 29; Mismatches 40; Indels 4; Gaps 4;

```

QY 1 VEKLMVTVVYGVPRWKEATTTTFCASDAKAYDEVHNVMAHACVPTDPNPQEVLENVT 60
DB 25 VEKLMVTVVYGVPRWKEATTTTFCASDAKAYDEVHNVMAHACVPTDPNPQEVLENVT 84
QY 61 EHFNMKNKNWQMOEDIIISLMDOSLRPCVKLTPLCVTLNCKDV-NATNTTNSDEGTMER 119
DB 85 EHFNMKNKNWQMOEDIIISLMDOSLRPCVKLTPLCVTLNCKDV-NATNTTNSDEGTMER 144
QY 120 GEIKNCSFNITTSIRDEVQKEYALFYKLDVVPIDNNNTSRLISCDTSVITQACPKISFE 179
DB 145 GEIKNCSFKVTTTSIRNMQKEYALFYKLDVVPIDNNNTSYKLINCNTSVITQACPKISFE 204
QY 180 PIPHYCAPAGFALIKCNDKTFNGKPCCKVSTVOCTHGIRPVVSTOLLNGLSLAESEV 239
DB 205 PIPHYCAPAGFALIKCNDKTFNGKPCCKVSTVOCTHGIRPVVSTOLLNGLSLAESEV 264
QY 240 IRSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGRAFYTTGIIIGDIRQAHN 299
DB 265 IRSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGRAFYTTGIIIGDIRQAHN 324
QY 300 ISRAKMDTLKQIYIKRBEQENKTIYFNHSSGDEPEIVHSHFNCGEFFYCNSDTOLFNS 359
DB 325 ISRAKMDTLKQIYIKRBEQENKTIYFNHSSGDEPEIVHSHFNCGEFFYCNSDTOLFNS 384
QY 360 TANNNTSGSNNTGNTTTLPCRICKOIIINMOEYKAMVAPIRGOIRCSSNITGLLTRD 419
DB 385 TW-NNTIGPNNTNG-TITLPCRICKOIIINMOEYKAMVAPIRGOIRCSSNITGLLTRD 442
QY 420 GGIN-ENGTEIFRPGGDMRDNMRSEFYKYKVKVIEPLGVAFTKAKRRVVQREKRAVIG 478
DB 443 GGEISNTTEIFRPGGDMRDNMRSEFYKYKVKVIEPLGVAFTKAKRRVVQREKRAVIG 502
QY 479 AVPLGFIGAAGSTMGASMTLTVQARLLSGIVQOQNNLLRAIEAOQRLQITWGIKOL 538
DB 503 AMPLGFIGAAGSTMGASMTLTVQARLLSGIVQOQNNLLRAIEAOQRLQITWGIKOL 562
QY 539 QARLAVERYLKQOQLIGWCSGKLTCTTAVPNMASWSKSLDRIMNNMTWMEWEID 598
DB 563 QARLAVERYLKQOQLIGWCSGKLTCTTAVPNMASWSKSLDRIMNNMTWMEWEID 622
QY 599 NYTSEIYTLIEESQNOQEKNEQELLEDKMASLMMWPDITNMLWY 643
DB 623 NYTSEIYTLIEESQNOQEKNEQELLEDKMASLMMWPDITNMLWY 667

```

```

RESULT 6
US-190-434B-2
; Sequence 2, Application US/10190434B
; Publication No. US20030194800A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B
; FILE REFERENCE: 2300-1621.20
; CURRENT APPLICATION NUMBER: US/10/190,434B
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 842
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SF162
US-10-190-434B-2

```

Query Match 87.8%; Score 3041; DB 12; Length 842;
 Best Local Similarity 88.7%; Pred. No. 9.6e-270;
 Matches 572; Conservative 29; Mismatches 40; Indels 4; Gaps 4;

```

QY 1 VEKLMVTVVYGVPRWKEATTTTFCASDAKAYDEVHNVMAHACVPTDPNPQEVLENVT 60
DB 25 VEKLMVTVVYGVPRWKEATTTTFCASDAKAYDEVHNVMAHACVPTDPNPQEVLENVT 84
QY 61 EHFNMKNKNWQMOEDIIISLMDOSLRPCVKLTPLCVTLNCKDV-NATNTTNSDEGTMER 119
DB 85 EHFNMKNKNWQMOEDIIISLMDOSLRPCVKLTPLCVTLNCKDV-NATNTTNSDEGTMER 144
QY 120 GEIKNCSFNITTSIRDEVQKEYALFYKLDVVPIDNNNTSRLISCDTSVITQACPKISFE 179
DB 145 GEIKNCSFKVTTTSIRNMQKEYALFYKLDVVPIDNNNTSYKLINCNTSVITQACPKISFE 204
QY 180 PIPHYCAPAGFALIKCNDKTFNGKPCCKVSTVOCTHGIRPVVSTOLLNGLSLAESEV 239
DB 205 PIPHYCAPAGFALIKCNDKTFNGKPCCKVSTVOCTHGIRPVVSTOLLNGLSLAESEV 264
QY 240 IRSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGRAFYTTGIIIGDIRQAHN 299
DB 265 IRSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGRAFYTTGIIIGDIRQAHN 324
QY 300 ISRAKMDTLKQIYIKRBEQENKTIYFNHSSGDEPEIVHSHFNCGEFFYCNSDTOLFNS 359
DB 325 ISRAKMDTLKQIYIKRBEQENKTIYFNHSSGDEPEIVHSHFNCGEFFYCNSDTOLFNS 384
QY 360 TANNNTSGSNNTGNTTTLPCRICKOIIINMOEYKAMVAPIRGOIRCSSNITGLLTRD 419
DB 385 TW-NNTIGPNNTNG-TITLPCRICKOIIINMOEYKAMVAPIRGOIRCSSNITGLLTRD 442
QY 420 GGIN-ENGTEIFRPGGDMRDNMRSEFYKYKVKVIEPLGVAFTKAKRRVVQREKRAVIG 478
DB 443 GGEISNTTEIFRPGGDMRDNMRSEFYKYKVKVIEPLGVAFTKAKRRVVQREKRAVIG 502
QY 479 AVPLGFIGAAGSTMGASMTLTVQARLLSGIVQOQNNLLRAIEAOQRLQITWGIKOL 538
DB 503 AMPLGFIGAAGSTMGASMTLTVQARLLSGIVQOQNNLLRAIEAOQRLQITWGIKOL 562
QY 539 QARLAVERYLKQOQLIGWCSGKLTCTTAVPNMASWSKSLDRIMNNMTWMEWEID 598
DB 563 QARLAVERYLKQOQLIGWCSGKLTCTTAVPNMASWSKSLDRIMNNMTWMEWEID 622
QY 599 NYTSEIYTLIEESQNOQEKNEQELLEDKMASLMMWPDITNMLWY 643
DB 623 NYTSEIYTLIEESQNOQEKNEQELLEDKMASLMMWPDITNMLWY 667

```

```

RESULT 7
US-10-190-305A-2
; Sequence 2, Application US/10190305A
; Publication No. US20030198621A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR
; FILE REFERENCE: 2302-18702 / 18702.002
; CURRENT APPLICATION NUMBER: US/10/190,305A
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 842
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SF162
US-10-190-305A-2

```

Query Match 87.8%; Score 3041; DB 12; Length 842;
 Best Local Similarity 88.7%; Pred. No. 9.6e-270;
 Matches 572; Conservative 29; Mismatches 40; Indels 4; Gaps 4;

```

Qy 1 VEKLVVTVYGVVWKEATTTLCASDAKAYTEVHNWATHACVPTDNPQEVILENT 60
Db 25 VEKLVVTVYGVVWKEATTTLCASDAKAYTEVHNWATHACVPTDNPQEVILENT 84
Qy 61 EHFNNKNNWVQWQMEDIIISLMDOSLKPCVLTPLCVTLNCKDV-NATNTDSEGTHER 119
Db 85 EHFNNKNNWVQWQMEDIIISLMDOSLKPCVLTPLCVTLNCKDV-NATNTDSEGTHER 144
Qy 120 GEIKNSCNITTSIRDEVOKEALFYKLDVVPIDNNNTSYRLISCDTSVITQACPKISFE 179
Db 145 GEIKNSCNITTSIRDEVOKEALFYKLDVVPIDNNNTSYRLISCDTSVITQACPKISFE 204
Qy 180 PIPHYCAPAGFAILKCDKDTFNGKPCKNVSTVOCTHGIRPVYSTOLLNGLSLEBEV 239
Db 205 PIPHYCAPAGFAILKCDKDTFNGKPCKNVSTVOCTHGIRPVYSTOLLNGLSLEBEV 264
Qy 240 IRSDFNTNNAKTIIVQLKESVEINCTRPNNTRKSIHIGPRAFYTTGIIIDIRQACHN 299
Db 265 IRSDFNTNNAKTIIVQLKESVEINCTRPNNTRKSIHIGPRAFYTTGIIIDIRQACHN 324
Qy 300 ISRAKMDTLKQIYVTLKREOFENKTIIVNHSFGGDEIYVHSGGDEFFYCNSTOLFNS 359
Db 325 ISGEKNNNTLKQIYVTLKREOFENKTIIVNHSFGGDEIYVHSGGDEFFYCNSTOLFNS 384
Qy 360 TNNNTGSSNTEGNTITLPCRIOIIMMOEYGVKAMTAPPIRGQIRCSSNITGLLTRD 419
Db 385 TNNNTGSSNTEGNTITLPCRIOIIMMOEYGVKAMTAPPIRGQIRCSSNITGLLTRD 442
Qy 420 GGIN-ENGTEIFRPGGDMRDNRSEFYKYVVKIEPLGVAFTKCKRRVOREKAVIG 478
Db 443 GGEKISNTTEIFRPGGDMRDNRSEFYKYVVKIEPLGVAFTKCKRRVOREKAVIG 502
Qy 479 AVFLGFLGAGSTWGAASMTLTVQARLLSGIVQOONNLRAIEAQRMLOLTWGIKOL 538
Db 503 AMFLGFLGAGSTWGAASMTLTVQARLLSGIVQOONNLRAIEAQRMLOLTWGIKOL 562
Qy 539 QARVLAVERYLKDQQLLGIMGCSGKLCCTTAVPNNASNSKSLDRIMNNMTWMEWERID 598
Db 563 QARVLAVERYLKDQQLLGIMGCSGKLCCTTAVPNNASNSKSLDRIMNNMTWMEWERID 622
Qy 599 NYTSEIYTLIEESONQOEKNEQELLEDKMASLNNWPDITWMLWY 643
Db 623 NYTSEIYTLIEESONQOEKNEQELLEDKMASLNNWPDITWMLWY 667

```

RESULT 8
US-09-476-242-2
Sequence 2, Application US/09476242

Patent No. US2002014663A1
GENERAL INFORMATION:
APPLICANT: BARNETT, Susan
APPLICANT: MARTIN, Eric
TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES
FILE REFERENCE: 1605.002
CURRENT APPLICATION NUMBER: US/09/476,242
CURRENT FILING DATE: 1999-12-30
NUMBER OF SEQ ID NOS: 26
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2
LENGTH: 847
TYPE: PRT
ORGANISM: Human immunodeficiency virus
US-09-476-242-2

Query Match 87.8%; Score 3041; DB 10; Length 847;
Best Local Similarity 88.7%; Pred. No. 9.7e-270;
Matches 572; Conservative 29; Mismatches 40; Indels 4; Gaps 4;

Qy 1 VEKLVVTVYGVVWKEATTTLCASDAKAYTEVHNWATHACVPTDNPQEVILENT 60
Db 30 VEKLVVTVYGVVWKEATTTLCASDAKAYTEVHNWATHACVPTDNPQEVILENT 89

```

Qy 61 EHFNNKNNWVQWQMEDIIISLMDOSLKPCVLTPLCVTLNCKDV-NATNTDSEGTHER 119
Db 90 EHFNNKNNWVQWQMEDIIISLMDOSLKPCVLTPLCVTLNCKDV-NATNTDSEGTHER 149
Qy 120 GEIKNSCNITTSIRDEVOKEALFYKLDVVPIDNNNTSYRLISCDTSVITQACPKISFE 179
Db 150 GEIKNSCNITTSIRDEVOKEALFYKLDVVPIDNNNTSYRLISCDTSVITQACPKISFE 209
Qy 180 PIPHYCAPAGFAILKCDKDTFNGKPCKNVSTVOCTHGIRPVYSTOLLNGLSLEBEV 239
Db 210 PIPHYCAPAGFAILKCDKDTFNGKPCKNVSTVOCTHGIRPVYSTOLLNGLSLEBEV 269
Qy 240 IRSDFNTNNAKTIIVQLKESVEINCTRPNNTRKSIHIGPRAFYTTGIIIDIRQACHN 299
Db 270 IRSDFNTNNAKTIIVQLKESVEINCTRPNNTRKSIHIGPRAFYTTGIIIDIRQACHN 329
Qy 300 ISRAKMDTLKQIYVTLKREOFENKTIIVNHSFGGDEIYVHSGGDEFFYCNSTOLFNS 359
Db 330 ISGEKNNNTLKQIYVTLKREOFENKTIIVNHSFGGDEIYVHSGGDEFFYCNSTOLFNS 389
Qy 360 TNNNTGSSNTEGNTITLPCRIOIIMMOEYGVKAMTAPPIRGQIRCSSNITGLLTRD 419
Db 390 TNNNTGSSNTEGNTITLPCRIOIIMMOEYGVKAMTAPPIRGQIRCSSNITGLLTRD 447
Qy 420 GGIN-ENGTEIFRPGGDMRDNRSEFYKYVVKIEPLGVAFTKCKRRVOREKAVIG 478
Db 448 GGEKISNTTEIFRPGGDMRDNRSEFYKYVVKIEPLGVAFTKCKRRVOREKAVIG 507
Qy 479 AVFLGFLGAGSTWGAASMTLTVQARLLSGIVQOONNLRAIEAQRMLOLTWGIKOL 538
Db 508 AMFLGFLGAGSTWGAASMTLTVQARLLSGIVQOONNLRAIEAQRMLOLTWGIKOL 567
Qy 539 QARVLAVERYLKDQQLLGIMGCSGKLCCTTAVPNNASNSKSLDRIMNNMTWMEWERID 598
Db 568 QARVLAVERYLKDQQLLGIMGCSGKLCCTTAVPNNASNSKSLDRIMNNMTWMEWERID 627
Qy 599 NYTSEIYTLIEESONQOEKNEQELLEDKMASLNNWPDITWMLWY 643
Db 628 NYTSEIYTLIEESONQOEKNEQELLEDKMASLNNWPDITWMLWY 672

```

RESULT 9
US-10-026-741-103
Sequence 103, Application US/10026741
Publication No. US2003004960A1
GENERAL INFORMATION:
APPLICANT: CHARNEAU, PIERRE
CLAVEL, FRANCOISE
BORMAN, ANDREW
OUILLENT, CAROLINE
GUETARD, DENISE
MONTAGNIER, LUC
DONJON DE SAINT-MARTIN, JACQUELINE
COHEN, JACQUES
SUBTYPE ANTIGENS
TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF HIV-1 TYPE (OR
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSSEE: Flinnegan, Henderson, Farabow, Garrett &
Dunner, L.L.P.
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/026,741
FILING DATE: 27-Dec-2001

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: US 08/817,441
 FILING DATE: 31-AUG-1998
 APPLICATION NUMBER: PCT/FR 95/01391
 FILING DATE: 20-OCT-1995
 APPLICATION NUMBER: FR 9412554
 FILING DATE: 20-OCT-1994
 APPLICATION NUMBER: FR 9502526
 FILING DATE: 03-MAR-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Meyers, Kenneth J.
 REGISTRATION NUMBER: 25,146
 REFERENCE/DOCKET NUMBER: 03260, 6005-00000
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-408-4000
 TELEFAX: 202-408-4400
 INFORMATION FOR SEQ ID NO: 103:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 861 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 103:
 US-10-026-741-103

Query Match 86.4%; Score 2992.5; DB 15; Length 861;
 Best Local Similarity 86.7%; Pred. No. 2.8e-265; Indels 15; Gaps 8;
 Matches 569; Conservative 32; Mismatches 40;

2 EKLAVTVYGVVWPKKATTTLLFCASDAKAYDEVHNVWATHACVPTDPNPOEVLEAVTE 61
 32 EKLAVTVYGVVWPKKATTTLLFCASDAKAYDEVHNVWATHACVPTDPNPOEVLEAVTE 91
 62 HFNMMKNMVEQWQEDIIISLMDQSLKPCVKLTPLCVTLNCKOV-NANT---TNDSEG- 115
 92 NFNMMKNMVEQWQEDIIISLMDQSLKPCVKLTPLCVSLKCTDLGNANTNNTSSSGE 151
 116 -TMRGEIKKNSFNITTSIRDEVQKEYALFYKLDVVPIDNNNTSRLISCDTSVITQACP 174
 152 MMEKGEIKKNSFNITTSIRDEVQKEYALFYKLDIIPIDNDTTSYKLSCTNSVTITQACP 211
 175 KISFPIPIHYCAPAGFALILKNDKTFNGKPCCKNVSTVQCTHGIRPVVSTOLLNGSLA 234
 212 KVSFPIPIHYCAPAGFALILKNDKTFNGKPCCKNVSTVQCTHGIRPVVSTOLLNGSLA 271
 235 EEEVVISRANFTDPAKTIIVQLNOSVEINCTRPNNNTKRSIHI--GPGAFYTTGEIIGD 292
 272 EEEVVISRANFTDPAKTIIVQLNOSVEINCTRPNNNTKRSIHI--GPGAFYTTGEIIGD 330
 293 IROHCHNISRAKMDTLKQIVIKLEQF-EKTIIVFNHSSGDEPVIWHSFNGCEFFYC 351
 331 MROHCHNISRAKMDTLKQIVIKLEQF-EKTIIVFNHSSGDEPVIWHSFNGCEFFYC 390
 352 NSTOLFSTWNN--NTEGSNNTEG-NTITLPCRILKQIINMMQEVGKAMVAPPIRGQIR 407
 391 NSTOLFSTWNN--NTEGSNNTEG-NTITLPCRILKQIINMMQEVGKAMVAPPIRGQIR 450
 408 SSNTITGLLTRDGINENGTETIFRPGGDMRDMRSEFYKYKVIKLEPLGVAFTCKRRV 467
 451 SSNTITGLLTRDGINENGTETIFRPGGDMRDMRSEFYKYKVIKLEPLGVAFTCKRRV 510
 468 VQREKRAVGTGAVLGLFGAAGSTMGASMTLTVQARLLSGIVQOONLLRAIEAOQRL 527
 511 VQREKRAVGTGAVLGLFGAAGSTMGASMTLTVQARLLSGIVQOONLLRAIEAOQRL 570
 528 LQITVWGIKQOLAVERYLGDQOLGIGWCSGKLICTTAVPNASMSKSLDRINWMT 587
 571 LQITVWGIKQOLAVERYLGDQOLGIGWCSGKLICTTAVPNASMSKSLDRINWMT 630
 588 MTWMEEREIDNTSEIYTLIEESQNOQEKNOBELLELDKVASLWNPEDITWMLY 643
 631 MTWMEEREIDNTSEIYTLIEESQNOQEKNOBELLELDKVASLWNPEDITWMLY 686

RESULT 10
US-09-476-242-1

Sequence 1, Application US/09476242
 Patent No. US2002014663A1
 GENERAL INFORMATION:
 APPLICANT: BARRETT, Susan
 APPLICANT: HARTOG, Karin
 APPLICANT: MARTIN, Eric
 TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES
 FILE REFERENCE: 1605,002
 CURRENT APPLICATION NUMBER: US/09/476,242
 CURRENT FILING DATE: 1999-12-30
 NUMBER OF SEQ ID NOS: 26
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 1
 LENGTH: 856
 TYPE: PRT
 ORGANISM: Human immunodeficiency virus
 US-09-476-242-1

Query Match 85.6%; Score 2964; DB 10; Length 856;
 Best Local Similarity 86.0%; Pred. No. 1.1e-262; Indels 10; Gaps 6;
 Matches 560; Conservative 36; Mismatches 45;

2 EKLAVTVYGVVWPKKATTTLLFCASDAKAYDEVHNVWATHACVPTDPNPOEVLEAVTE 61
 32 EKLAVTVYGVVWPKKATTTLLFCASDAKAYDEVHNVWATHACVPTDPNPOEVLEAVTE 91
 62 HFNMMKNMVEQWQEDIIISLMDQSLKPCVKLTPLCVTLNCKOV-NANT---TNDSEG- 119
 92 NFNMMKNMVEQWQEDIIISLMDQSLKPCVKLTPLCVSLKCTDLGNANTNNTSSSGE 151
 120 GEIKNSFNITTSIRDEVQKEYALFYKLDVVPIDNNNTSRLISCDTSVITQACP 179
 152 GEIKNSFNITTSIRDEVQKEYALFYKLDIIPIDNDTTSYKLSCTNSVTITQACP 211
 180 PIPHYCAPAGFALILKNDKTFNGKPCCKNVSTVQCTHGIRPVVSTOLLNGSLAEEV 239
 212 PIPHYCAPAGFALILKNDKTFNGKPCCKNVSTVQCTHGIRPVVSTOLLNGSLAEEV 271
 240 IRSDFNTNNAKTIIVQLNOSVEINCTRPNNNTKRSIHI--GPGAFYTTGEIIGDROAH 297
 272 IRSDFNTNNAKTIIVQLNOSVEINCTRPNNNTKRSIHI--GPGAFYTTGEIIGDROAH 330
 298 CNISRAKMDTLKQIVIKLEQF-EKTIIVFNHSSGDEPVIWHSFNGCEFFYCSTOL 356
 331 CNISRAKMDTLKQIVIKLEQF-EKTIIVFNHSSGDEPVIWHSFNGCEFFYCSTOL 390
 357 FNSTWNN--NTEGSNNTEG-NTITLPCRILKQIINMMQEVGKAMVAPPIRGQIR 412
 391 FNSTWNN--NTEGSNNTEG-NTITLPCRILKQIINMMQEVGKAMVAPPIRGQIR 450
 413 GLULTRDGINENGTETIFRPGGDMRDMRSEFYKYKVIKLEPLGVAFTCKRRV 472
 451 GLULTRDGINENGTETIFRPGGDMRDMRSEFYKYKVIKLEPLGVAFTCKRRV 510
 473 RANGIGAVTIGFGLAAGSTMGASMTLTVQARLLSGIVQOONLLRAIEAOQRL 532
 511 RANGIGAVTIGFGLAAGSTMGASMTLTVQARLLSGIVQOONLLRAIEAOQRL 570
 533 WGIKQOLAVERYLGDQOLGIGWCSGKLICTTAVPNASMSKSLDRINWMT 592
 571 WGIKQOLAVERYLGDQOLGIGWCSGKLICTTAVPNASMSKSLDRINWMT 630
 593 WEREIDNTSEIYTLIEESQNOQEKNOBELLELDKVASLWNPEDITWMLY 643
 631 WEREIDNTSEIYTLIEESQNOQEKNOBELLELDKVASLWNPEDITWMLY 681

RESULT 11
 US-10-196-515-11


```

; Sequence 11, Application US/10196515
; Publication No. US20030091594A1
; GENERAL INFORMATION:
; APPLICANT: LABRANCHE, James A.
; APPLICANT: HOBIE, James A.
; APPLICANT: LABRANCHE, Celia C.
; APPLICANT: DONS, Robert W.
; APPLICANT: HOFFMAN, Trevor L.
; TITLE OF INVENTION: CD4-INDEPENDENT HIV ENVELOPE PROTEINS AS VACCINES AND
; FILE REFERENCE: Hoxie 9596-104U1 (0282)
; CURRENT APPLICATION NUMBER: US/10/196,515
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: US/09/337,387
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: US 09/317,556
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 11
; LENGTH: 856
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-10-196-515-11

```

```

Query Match      85.3%; Score 2953; DB 15; Length 856;
Best Local Similarity 85.9%; Pred. No. 1.2e-261;
Matches 559; Conservative 36; Mismatches 46; Indels 10; Gaps 6;

QY 2 EKLWTVVYGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTE 61
DB 32 EKLWTVVYGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTE 91
QY 62 HFNMMKNMVEQMEDIIISLDQSLKPCVKLTPLCVTLNCKDVNATNTDSEG--TWER 119
DB 92 NFDMMKNDMEQMEDIIISLDQSLKPCVKLTPLCVSLKCTDLKNDVTNTSSSGRMIMEK 151
QY 120 GEIKNCSFNITTSIRDEYQKEYALFYKLDVVPIDNNNTSYLLISCDTSVITQACPKISFE 179
DB 152 GEIKNCSFNITTSIRDEYQKEYALFYKLDIIPIDNDTTSYLLISCDTSVITQACPKISFE 211
QY 180 PIPHYCAPAGFALLKNDKTFNGKGPCKNVSTVOCTHGIRPVVSTOLLNGLSLAEEV 239
DB 212 PIPHYCAPAGFALLKNDKTFNGKGPCKNVSTVOCTHGIRPVVSTOLLNGLSLAEEV 271
QY 240 IRSDFNNAKTIIVQLKESYEINCTRPNNNRKSIH--GPGRAFYTTGELIGDIRQAH 297
DB 272 IRSDFNNAKTIIVQLKESYEINCTRPNNNRKSIH--GPGRAFYTTGELIGDIRQAH 330
QY 298 CNISRAKKNNDTLKQIVIKLREOF--ENKTIIVFNHSSGDPBEIVMHSFNCEGEFFYCNSTOL 356
DB 331 CNISRAKKNNDTLKQIVIKLREOF--ENKTIIVFNHSSGDPBEIVMHSFNCEGEFFYCNSTOL 390
QY 357 FNSTNN--NTEGSNNTGEG--NTITLPCRIKQIIMMOEYKAMAYAPPIRQOICSSNITL 412
DB 391 FNSTNN--NTEGSNNTGEG--NTITLPCRIKQIIMMOEYKAMAYAPPIRQOICSSNITL 450
QY 413 GLLTRDGINENGEIIFRPGGDMRDNRSEFYKYVVKIIEPLGVAPTKCRVRVOREK 472
DB 451 GLLTRDGINENGEIIFRPGGDMRDNRSEFYKYVVKIIEPLGVAPTKCRVRVOREK 510
QY 473 RAVGIGAVFLGFLGAAGSTMGASMTLTVQARLLLSGIVOOQNNLLRAIEAQOQMLQTLV 532
DB 511 RAVGIGAVFLGFLGAAGSTMGASMTLTVQARLLLSGIVOOQNNLLRAIEAQOQMLQTLV 570
QY 533 WGIQOLARVLAVERYLGDQOLGIGWCSGKLICTTAVPNNASNSKSLDRIMNMTTME 592
DB 571 WGIQOLARVLAVERYLGDQOLGIGWCSGKLICTTAVPNNASNSKSLDRIMNMTTME 630
QY 593 WEREIDNTSEIYTLIEESQNOEKNQOELELDKMASLNMWPDITNMLWY 643
DB 631 WEREIDNTSEIYTLIEESQNOEKNQOELELDKMASLNMWPDITNMLWY 681

```

```

RESULT 12
US-10-196-515-3
; Sequence 3, Application US/10196515
; Publication No. US20030091594A1
; GENERAL INFORMATION:
; APPLICANT: LABRANCHE, James A.
; APPLICANT: HOBIE, James A.
; APPLICANT: LABRANCHE, Celia C.
; APPLICANT: DONS, Robert W.
; APPLICANT: HOFFMAN, Trevor L.
; TITLE OF INVENTION: CD4-INDEPENDENT HIV ENVELOPE PROTEINS AS VACCINES AND
; FILE REFERENCE: Hoxie 9596-104U1 (0282)
; CURRENT APPLICATION NUMBER: US/10/196,515
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: US/09/337,387
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: US 09/317,556
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3
; LENGTH: 726
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-10-196-515-3

```

```

Query Match      84.1%; Score 2910.5; DB 15; Length 726;
Best Local Similarity 84.7%; Pred. No. 7.2e-258;
Matches 549; Conservative 39; Mismatches 51; Indels 9; Gaps 6;

QY 2 EKLWTVVYGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTE 61
DB 32 EKLWTVVYGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTE 91
QY 62 HFNMMKNMVEQMEDIIISLDQSLKPCVKLTPLCVTLNCKDVNATNTDSEG--TWER 119
DB 92 NFNMMKNDMEQMEDIIISLDQSLKPCVKLTPLCVSLKCTDLKNDVTNTSSSGRMIMEK 151
QY 120 GEIKNCSFNITTSIRDEYQKEYALFYKLDVVPIDNNNTSYLLISCDTSVITQACPKISFE 179
DB 152 GEIKNCSFNITTSIRDEYQKEYALFYKLDIIPIDNDTTSYLLISCDTSVITQACPKISFE 211
QY 180 PIPHYCAPAGFALLKNDKTFNGKGPCKNVSTVOCTHGIRPVVSTOLLNGLSLAEEV 239
DB 212 PIPHYCAPAGFALLKNDKTFNGKGPCKNVSTVOCTHGIRPVVSTOLLNGLSLAEEV 271
QY 240 IRSDFNNAKTIIVQLKESYEINCTRPNNNRK--SIHIGPRAFYTTGELIGDIRQAH 297
DB 272 IRSDFNNAKTIIVQLKESYEINCTRPNNNRK--SIHIGPRAFYTTGELIGDIRQAH 330
QY 298 CNISRAKKNNDTLKQIVIKLREOF--ENKTIIVFNHSSGDPBEIVMHSFNCEGEFFYCNSTOL 356
DB 331 CNISRAKKNNDTLKQIVIKLREOF--ENKTIIVFNHSSGDPBEIVMHSFNCEGEFFYCNSTOL 390
QY 357 FNSTNN--NTEGSNNTGEG--NTITLPCRIKQIIMMOEYKAMAYAPPIRQOICSSNITL 415
DB 391 FNSTNN--NTEGSNNTGEG--NTITLPCRIKQIIMMOEYKAMAYAPPIRQOICSSNITL 448
QY 416 LTRDGINENGEIIFRPGGDMRDNRSEFYKYVVKIIEPLGVAPTKCRVRVOREK 475
DB 449 LTRDGINENGEIIFRPGGDMRDNRSEFYKYVVKIIEPLGVAPTKCRVRVOREK 508
QY 476 GIGAVFLGFLGAAGSTMGASMTLTVQARLLLSGIVOOQNNLLRAIEAQOQMLQTLV 535
DB 509 GIGAVFLGFLGAAGSTMGASMTLTVQARLLLSGIVOOQNNLLRAIEAQOQMLQTLV 568
QY 536 KQOLARVLAVERYLGDQOLGIGWCSGKLICTTAVPNNASNSKSLDRIMNMTTME 595
DB 569 KQOLARVLAVERYLGDQOLGIGWCSGKLICTTAVPNNASNSKSLDRIMNMTTME 628
QY 596 EIDNTSEIYTLIEESQNOEKNQOELELDKMASLNMWPDITNMLWY 643
DB 629 EIDNTSEIYTLIEESQNOEKNQOELELDKMASLNMWPDITNMLWY 676

```

RESULT 13
US-10-196-515-12
; Sequence 12, Application US/10196515
; Publication No. US20030091594A1
; GENERAL INFORMATION:
; APPLICANT: HOXIE, James A.
; APPLICANT: LABRANCHE, Celia C.
; APPLICANT: DOMS, Robert W.
; APPLICANT: HOFFMAN, Trevor L.
; TITLE OF INVENTION: CD4-INDEPENDENT HIV ENVELOPE PROTEINS AS VACCINES AND
; TITLE OF INVENTION: THERAPEUTICS
; FILE REFERENCE: Hoxie 9596-104U1 (0282)
; CURRENT APPLICATION NUMBER: US/10/196,515
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: US/09/337,387
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: US 09/317,556
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 12
; LENGTH: 759
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-10-196-515-12

Query Match 83.7%; Score 2897.5; DB 15; Length 759;
Best Local Similarity 84.6%; Pred. No. 1,2e-256;
Matches 548; Conservative 40; Mismatches 51; Indels 9; Gaps 6;
QY 2 EKLWTVYGVVWKEATTTLFCASDAKAYDEVHNVNATACVPTDNPQEVLENTENF 61
DB 32 EKLWTVYGVVWKEATTTLFCASDAKAYDEVHNVNATACVPTDNPQEVLENTENF 91
QY 62 HFNWKNWVBOHEDIIISLMDQSLKPCVKLTPLCVTLNCKDVNATNTNDSSEG--TWER 119
DB 92 NFNWKNWVBOHEDIIISLMDQSLKPCVKLTPLCVTLNCKDVNATNTNDSSEGSRMIMEX 151
QY 120 GEIKNGSFNITTSIRDEVOKEYALFYKLDVVPIDNNNTSYRLISCDTSVITQACPISFE 179
DB 152 GEIKNGSFNITTSIRDEVOKEYALFYKLDVVPIDNNNTSYRLISCDTSVITQACPISFE 211
QY 180 PIRPHYCAPGFAFKNDKTFNGKPCKNVSTVOCTHGRIPVYSTOLLNGSLAEERYV 239
DB 212 PIRPHYCAPGFAFKNDKTFNGKPCKNVSTVOCTHGRIPVYSTOLLNGSLAEERYV 271
QY 240 IRSDFNNAKTIIVQLKESVEINCTRPNNNTRKSIH--GPGRAFTYTGIIIGDIRQAH 297
DB 272 IRSDFNNAKTIIVQLKESVEINCTRPNNNTRKSIH--GPGRAFTYTGIIIGDIRQAH 330
QY 298 CNISRAKMWDTLKOIVIKLREOF--ENKTIYFNHSSGGDPETVMSFNGCEGFYCNSTOL 356
DB 331 CNISRAKMWDTLKOIVIKLREOF--ENKTIYFNHSSGGDPETVMSFNGCEGFYCNSTOL 390
QY 357 FNSWNNNTBESNNTG--NTTLPCRITKOIINMOEVKAMYPAPPIGQIRCSNITGL 415
DB 391 FNSWNNNTBESNNTG--NTTLPCRITKOIINMOEVKAMYPAPPIGQIRCSNITGL 448
QY 416 LTRGGINENGTEIFRPGGDMRDMRSEFYKVKVLEPLGVAPTCKRRVVOREKRAV 475
DB 449 LTRGGINENGTEIFRPGGDMRDMRSEFYKVKVLEPLGVAPTCKRRVVOREKRAV 508
QY 476 GIGAVPLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOQNNLLRAIEAQHMLQITWGI 535
DB 509 GIGAVPLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOQNNLLRAIEAQHMLQITWGI 568
QY 536 KQLOARVLAVERYLKDOOLGIGWCSGKLICTTAVPWNASWSNLSLDRINNNMTWEMER 595
DB 569 KQLOARVLAVERYLKDOOLGIGWCSGKLICTTAVPWNASWSNLSLDRINNNMTWEMER 628
QY 596 EIDNYTSEIYTLIESQNOQEKNEQELLELDKMASLWNVWFDITWLVWY 643

DB 629 EINNYSLSHLSIESQNOQEKNEQELLELDKMASLWNVWFDITWLVWY 676
RESULT 14
US-09-938-406-1
; Sequence 1, Application US/09938406
; Patent No. US20020155120A1
; GENERAL INFORMATION:
; APPLICANT: Lowell, George
; APPLICANT: Vancolt, Thomas
; APPLICANT: Birt, Deborah
; TITLE OF INVENTION: PROTEIN AND PEPTIDE VACCINES FOR
; TITLE OF INVENTION: INDUCING MUCOSAL IMMUNITY
; FILE REFERENCE: 40646-20002.10
; CURRENT APPLICATION NUMBER: US/09/938,406
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 09/214,701
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/US 97/12253
; PRIOR FILING DATE: 1997-07-10
; PRIOR APPLICATION NUMBER: US 60/021,687
; PRIOR FILING DATE: 1996-07-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 868
; TYPE: PRT
; ORGANISM: Virus HIV-1
US-09-938-406-1

Query Match 83.6%; Score 2895.5; DB 10; Length 868;
Best Local Similarity 82.1%; Pred. No. 2,2e-256;
Matches 541; Conservative 40; Mismatches 59; Indels 19; Gaps 4;
QY 4 LMTVYGVVWKEATTTLFCASDAKAYDEVHNVNATACVPTDNPQEVLENTENF 63
DB 35 LMTVYGVVWKEATTTLFCASDAKAYDEVHNVNATACVPTDNPQEVLENTENF 94
QY 64 NFNWKNWVBOHEDIIISLMDQSLKPCVKLTPLCVTLNCKDVNATNTNDSSEGTM----- 117
DB 95 NFNWKNWVBOHEDIIISLMDQSLKPCVKLTPLCVTLNCKDVNATNTNDSSEGTM----- 154
QY 118 -BERGEIKNGSFNITTSIRDEVOKEYALFYKLDVVPIDN-----NNTSYRLISCDTSVITQ 171
DB 155 -BERGEIKNGSFNITTSIRDEVOKEYALFYKLDVVPIDN-----NNTSYRLISCDTSVITQ 214
QY 172 ACPKISFEPIPIHYCAPGFAFKNDKTFNGKPCKNVSTVOCTHGRIPVYSTOLLNG 231
DB 215 ACPKISFEPIPIHYCAPGFAFKNDKTFNGKPCKNVSTVOCTHGRIPVYSTOLLNG 274
QY 232 SLAEEVYIISDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPGRAFTYTGIIIG 291
DB 275 SLAEEVYIISDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPGRAFTYTGIIIG 334
QY 292 DIRQAHNISRANKMDTLKOIVIKLREOFENKTIYFNHSSGGDPETVMSFNGCEGFYCN 351
DB 335 DIRQAHNISRANKMDTLKOIVIKLREOFENKTIYFNHSSGGDPETVMSFNGCEGFYCN 394
QY 352 NSTOLFNSWNNNTBESNNTG-----NTEGNTTLPCRITKOIINMOEVKAMYPAPPIGQIR 405
DB 395 NSTOLFNSWNNNTBESNNTG-----NTEGNTTLPCRITKOIINMOEVKAMYPAPPIGQIR 448
QY 406 RCSSNITGLLITRDGGINENGTEIFRPGGDMRDMRSEFYKVKVLEPLGVAPTCKRR 465
DB 445 RCSSNITGLLITRDGGINENGTEIFRPGGDMRDMRSEFYKVKVLEPLGVAPTCKRR 514
QY 466 RYVOREKRAV--IGAVPLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOQNNLLRAIEAQ 524
DB 515 RYVOREKRAV--IGAVPLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOQNNLLRAIEAQ 574
QY 525 QRMQLTVWGIKQLOARVLAVERYLKDOOLGIGWCSGKLICTTAVPWNASWSNLSLDR 584
DB 574 QRMQLTVWGIKQLOARVLAVERYLKDOOLGIGWCSGKLICTTAVPWNASWSNLSLDR 628

Db 575 OHLLQLTWGIKQLOARILAVERYLKDOQLLGFMCSSGKLICTTAVPNWNASNKTLDQI 634
Qy 585 WNNMTWMEWEREIDNYTSEIYTLIEBSONQOEKNEQELLELDKWSLNNWFDITNMLWY 643
Db 635 WNNMTWMEWEREIDNYTHLYTLIEBSONQOEKNEQELLELDKWSLNNWFDITNMLWY 693

RESULT 15
US-09-891-609-4
; Sequence 4, Application US/09891609
; Patent No. US20020127238A1
; GENERAL INFORMATION:
; APPLICANT: Stamatatos, Leonidas
; TITLE OF INVENTION: HIV-1 VACCINES AND SCREENING METHODS THEREFOR
; FILE REFERENCE: 2570-1-002N
; CURRENT APPLICATION NUMBER: US/09/891,609
; PRIOR FILING DATE: 2001-06-26
; PRIOR FILING DATE: 2000-06-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-09-891-609-4

Query Match 82.7%; Score 2862; DB 10; Length 619;

Best local similarity 84.5%; Pred.No.1.6e-253;

Matches 545; Conservative 25; Mismatches 43; Indels 32; Gaps 6;

Qy 1 VEKLVAVTVYGVVWKEATTTLLFCASDAKAYDTEVHNVATHACVPTDPNPOEVLENTY 60
Db 3 VEKLVAVTVYGVVWKEATTTLLFCASDAKAYDTEVHNVATHACVPTDPNPOEVLENTY 62
Qy 61 EHPNMMKNNWVEQWQEDIIISLWDQSLKPCVKLTPLCVTLNCKDV-NATNTNDSEGTWER 119
Db 63 ENFMNMKNNWVEQWQEDIIISLWDQSLKPCVKLTPLCVTLNCKDV-NATNTNDSEGTWER 122
Qy 120 GEIKKCSFNITTSIRDEVOKEALFYKLDVPRIDNNNTSYRLISCDTSVITQACPKISFE 179
Db 123 GEIKKCSFVVGAG-----KLNCTNSVITQACPKVSFE 155
Qy 180 PIPHYCAPAGFAILKCNDKTFNGKPCKNVSTVOCTHGIRPVSTOLLNGLAEEVY 239
Db 156 PIPHYCAPAGFAILKCNDKTFNGKPCKNVSTVOCTHGIRPVSTOLLNGLAEEVY 215
Qy 240 IRSNFTNNAKTIIVQLKESVEINCTRPNNNTKRSIHIGPRAFYTTGEIIGDIRQAHCN 299
Db 216 IRSENFNDNAKTIIVQLKESVEINCTRPNNNTKRSIHIGPRAFYTTGEIIGDIRQAHCN 275
Qy 300 ISRAKKNNTLKOIYVLAKEQFENKTIIVPHSSGDEPEIYMSFNCGEFFYCNSTOLFNS 359
Db 276 ISGERKNNNTLKOIYVLAKEQFENKTIIVPHSSGDEPEIYMSFNCGEFFYCNSTOLFNS 335
Qy 360 TWNNTEGSNNTEGNTITLPCRIKOIINWMOVGKAMYAPPIRGQIRCSNITGLLTRD 419
Db 336 TW-NNTIGPNNTNG-TITLPCRIKOIINWMOVGKAMYAPPIRGQIRCSNITGLLTRD 393
Qy 420 GGIN-ENGTEIFRPGGDMRDWRSEFYKYKVKIIEPLGVAPTKCRRVQREKRAVGIG 478
Db 394 GGKEISNTTEIFRPGGDMRDWRSELYKYKVKIIEPLGVAPTKARVRVQREKRAVTLG 453
Qy 479 AVFLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOQNNLRAIEAOQRMLOLTWGIKOL 538
Db 454 AMPLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOQNNLRAIEAOQRMLOLTWGIK-L 512
Qy 539 QARVLAVERYLGDQOLLGIWGCSSGLICTTAVPNWNASNKSILDRIMNNMTWMEWEREID 598
Db 513 QARVLAVERYLKDQOLLGIWGCSSGLICTTAVPNWNASNKSILDOIMNNMTWMEWEREID 572
Qy 599 NTSSEIYTLIEBSONQOEKNEQELLELDKWSLNNWFDITNMLWY 643
Db 573 NYTNLIYTLIEBSONQOEKNEQELLELDKWSLNNWFDISKMLWY 617

Db 573 NYTNLIYTLIEBSONQOEKNEQELLELDKWSLNNWFDISKMLWY 617
Search completed: December 12, 2003, 12:45:21
Job time : 28.8062 secs

THIS PAGE BLANK (US.70)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 12:27:29 ; Search time 14.9697 Seconds

(without alignments)
4130.780 Million cell updates/sec

Title: US-10-032-162-13

Perfect score: 3462

Sequence: 1 VEKLMVTYYGVVPWKEAT...ELDKWASIMNWDITNLMWY 643

Scoring table: BLOSUM62

Gapop 10.0 ; Gapext 0.5

Searched: 283308 seqs, 96168682 residues

1 number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3424	98.9	847	2 T09448	envelope glycoprot
2	3395	98.1	847	2 S13289	env protein - huma
3	3007.5	86.9	852	2 T12016	envelope glycoprot
4	3000	86.7	843	1 H44001	env polypeptide pr
5	2992.5	86.4	861	1 VCLJLV	env polypeptide pr
6	2989	86.3	852	1 VCLJBR	env polypeptide pr
7	2978	86.0	856	1 VCLJH3	env polypeptide pr
8	2974	85.9	854	1 VCLJVL	env polypeptide pr
9	2872	85.8	854	2 S13288	env polypeptide pr
10	2872	85.8	855	1 VCLJAJ	env protein - huma
11	2967.5	85.7	859	1 VCLJMN	env polypeptide pr
12	2967	85.7	861	1 VCLJSC	env polypeptide pr
13	2958.5	85.5	856	1 VCLJW	env polypeptide pr
14	2953.5	85.3	851	2 S33985	env polypeptide pr
15	2895.5	83.6	729	1 VCLJX	env polypeptide pr
16	2895.5	83.6	861	1 VCLJKB	env polypeptide pr
17	2895.5	83.6	868	1 VCLJH4	env polypeptide pr
18	2706	78.2	855	1 VCLJZR	env polypeptide pr
19	2686	77.6	853	2 S54384	envelope glycoprote
20	2657.5	76.8	859	2 T01672	envelope polypepte
21	2638.5	76.2	856	1 A44963	envelope polypepte
22	2614.5	75.3	846	1 VCLJND	env polypeptide pr
23	2244	64.8	854	2 A40218	env polypeptide pr
24	2165	62.5	506	2 A40218	envelop glycoprote
25	2002.5	57.8	443	2 C41621	env polypeptide M
26	1985.5	57.4	445	2 A41621	env polypeptide M
27	1938.5	56.0	495	2 S31493	env polypeptide M
28	1864	53.8	877	2 S49197	envelope protein p
29	1846	53.3	454	2 B41621	env polypeptide p

30	1825	52.7	863	2 A53034	gag polypeptide -
31	1320.5	38.1	290	2 S25940	env protein - huma
32	1299.5	37.5	297	2 S60538	envelope polypepte
33	1206	34.8	859	1 VCLJST	env polypeptide pr
34	1204.5	34.8	852	1 VCLJGS	env polypeptide pr
35	1201.5	34.7	299	2 S60554	envelope polypepte
36	1196.5	34.6	299	2 S60553	envelope polypepte
37	1191.5	34.4	299	2 S60552	envelope polypepte
38	1188	34.3	712	1 VCLJSA	env polypeptide pr
39	1187.5	34.3	299	2 S60551	envelope polypepte
40	1186.5	34.3	299	2 S60528	envelope polypepte
41	1186	34.3	294	2 S60545	envelope polypepte
42	1184	34.2	294	2 S60525	envelope polypepte
43	1183.5	34.2	299	2 S60529	envelope polypepte
44	1180.5	34.1	299	2 S60521	envelope polypepte
45	1180	34.1	877	2 C46356	env polypeptide -

ALIGNMENTS

RESULT 1

T09448
envelope glycoprotein - human immunodeficiency virus type 1 (strain JREF)

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 24-Nov-1999

C:Accession: T09448

R:Pang, S.; Vinters, H.V.; Akashi, T.; O'Brien, W.A.; Chen, I.S.; Koyanagi, Y.; Nemeroff, R.

Submitted to the EMBL Data Library, July 1996

A:Reference number: Z16673

A:Accession: T09448

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-847 <PAM>

A:Cross-references: EMBL:U63632; NID:G1465777; PID:G1465781

C:Genetics:

A:Gene: env

C:Superfamily: type B retrovirus env polypeptide

Query Match	98.9%	Score 3424	DB 2	Length 847
Best Local Similarity	99.2%	Pred. No. 4.2e-239		
Matches 638	Conservative 0	Mismatches 5	Indels 0	Gaps 0
QY	1	VEKLMVTYYGVVPWKEATTTLPASDAKADTEVHNMATHACVTPDPNPOEVLENT	60	
DB	30	VEKLMVTYYGVVPWKEATTTLPASDAKADTEVHNMATHACVTPDPNPOEVLENT	89	
QY	61	EHFMWKNVVEQWQEDIIISLMDQSLKPCVLTPLCVTLNCKDVNATNTNDSGTMERG	120	
DB	90	EHFMWKNVVEQWQEDIIISLMDQSLKPCVLTPLCVTLNCKDVNATNTNDSGTMERG	149	
QY	121	EIKKCSFRTTTSIDNVEKALFYKLDVVDINNNTSYRLISCDTSYTTACPKISFEP	180	
DB	150	EIKKCSFRTTTSIDNVEKALFYKLDVVDINNNTSYRLISCDTSYTTACPKISFEP	209	
QY	181	IPHYCAPAGFAIKCDKPTNGKPGCNVSTVCTGIRVSTTOLLNLSLAEVEVI	240	
DB	210	IPHYCAPAGFAIKCDKPTNGKPGCNVSTVCTGIRVSTTOLLNLSLAEVEVI	269	
QY	241	RSDNFTNNAKTIIVQLKESVEINCTRPNNTRKSIHIGPRAFYTTGHIIDIRQAHNI	300	
DB	270	RSDNFTNNAKTIIVQLKESVEINCTRPNNTRKSIHIGPRAFYTTGHIIDIRQAHNI	329	
QY	301	SRACNDTLKQIVIKLAEQFENKTIIVRHSGGPEIYMHSGFNGEFTYNSQLPFST	360	
DB	330	SRACNDTLKQIVIKLAEQFENKTIIVRHSGGPEIYMHSGFNGEFTYNSQLPFST	389	
QY	361	WNNTEGSNTGEGTITLPCRIKQIIMWQGVGKAMTAPPRGQIRCSNTTGLLTRDG	420	
DB	390	WNNTEGSNTGEGTITLPCRIKQIIMWQGVGKAMTAPPRGQIRCSNTTGLLTRDG	449	
QY	421	GINENGTEIFRPGGDMRWNRSEFYKXKVIPLGVAPTKCRRVQREKRAVGICAV	480	

Db 450 GINENGTEIFRPGGDMRDNRSELYKYKVKIEPLGVAPTKARRVOREKAVGIGAV 509
Qy 481 FLGFLGAAGSTMGAASMTLTVQARLLLSGIYVOQNNLLRAIEAQORMLQLTWNGIKQLOA 540
Db 510 FLGFLGAAGSTMGAASMTLTVQARLLLSGIYVOQNNLLRAIEAQORMLQLTWNGIKQLOA 569
Qy 541 RLAVERYLDGDOQLIGIWCSSGKLICTTAVPWNASWSNKSIDRIIMNMNTWMEEREIDNY 600
Db 570 RLAVERYLDGDOQLIGIWCSSGKLICTTAVPWNASWSNKSIDRIIMNMNTWMEEREIDNY 629
Qy 601 TSEIYTLIESQNOEKNOEQLLELDKMASLWNPFDITNNLMY 643
Db 630 TSEIYTLIESQNOEKNOEQLLELDKMASLWNPFDITNNLMY 672

RESULT 2

env protein - human immunodeficiency virus type 1
S13289
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997
Accession: S13289
R:O'Brien, W.A.; Koyanagi, Y.; Namazie, A.; Zhao, J.Q.; Diagne, A.; Idler, K.; Zack, J.A.
Nature 348, 63-73, 1990
A:Title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120
A:Reference number: S13289; PMID:91043044; PMID:2172833
A:Accession: S13289
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-847 <OBR>
C:Superfamily: type E retrovirus env polypeptide

Query Match 98.1%; Score 3395; DB 2; Length 847;
Best Local Similarity 98.6%; Pred. No. 5.2e-237;
Matches 633; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 2 EKLMTVYVYGVPMKEATTLTFCASDAKAYDEVHNVMAHACVPTDPNPOEVLNVTE 61
Db 31 EKLMTVYVYGVPMKEATTLTFCASDAKAYDEVHNVMAHACVPTDPNPOEVLNVTE 90
Qy 62 HENMKNMVEQNOEDIIISLDOSLKPCKVLTPLCVTLNCDVNAVNTTNDSEGMERGE 121
Db 91 HENMKNMVEQNOEDIIISLDOSLKPCKVLTPLCVTLNCDVNAVNTTNDSEGMERGE 150
Qy 122 IKKCSFNITTSIRDEVOKEYALFKLDVVPIDNNNTSYRLISCDTSVITQACPKISFEPI 181
Db 151 IKKCSFNITTSIRDEVOKEYALFKLDVVPIDNNNTSYRLISCDTSVITQACPKISFEPI 210
Qy 182 PIHYCAPAGFALIKCNDKTFNGKPCCKNVSTVQCTHGIRPVSTOLLNGLAEDEVVIR 241
Db 211 PIHYCAPAGFALIKCNDKTFNGKPCCKNVSTVQCTHGIRPVSTOLLNGLAEDEVVIR 270
Qy 242 SDNFTNNAKTIIVOLKESVEINCTRPNNNTKSHIIGRAFYTTGIIIGDIRAHGNIS 301
Db 271 SDNFTNNAKTIIVOLKESVEINCTRPNNNTKSHIIGRAFYTTGIIIGDIRAHGNIS 330
Qy 302 RAKKNDTLKOIVILRBOFEKTIIVFNHSSGDEPEIWNHSGGEFFYCNSITGLLFTSTW 361
Db 331 RAKKNDTLKOIVILRBOFEKTIIVFNHSSGDEPEIWNHSGGEFFYCNSITGLLFTSTW 390
Qy 362 NNNTGSGNNTGNTIILPCRIKOIINMOEYKAMVAPIRGOIRCSNITGLLFTDGG 421
Db 391 NNNTGSGNNTGNTIILPCRIKOIINMOEYKAMVAPIRGOIRCSNITGLLFTDGG 450
Qy 422 INENGTETFRPGGDMRDNRSELYKYKVKIEPLGVAPTKARRVOREKAVGIGAV 481
Db 451 INENGTETFRPGGDMRDNRSELYKYKVKIEPLGVAPTKARRVOREKAVGIGAV 510
Qy 482 LGFLGAAGSTMGAASMTLTVQARLLLSGIYVOQNNLLRAIEAQORMLQLTWNGIKQLOA 541
Db 511 LGFLGAAGSTMGAASMTLTVQARLLLSGIYVOQNNLLRAIEAQORMLQLTWNGIKQLOA 570
Qy 542 VLAVERYLDGDOQLIGIWCSSGKLICTTAVPWNASWSNKSIDRIIMNMNTWMEEREIDNY 601
Db 601 VLAVERYLDGDOQLIGIWCSSGKLICTTAVPWNASWSNKSIDRIIMNMNTWMEEREIDNY 672

Db 571 VLAVERYLDGDOQLIGIWCSSGKLICTTAVPWNASWSNKSIDRIIMNMNTWMEEREIDNY 630
Qy 602 SEIYTLIESQNOEKNOEQLLELDKMASLWNPFDITNNLMY 643
Db 631 SEIYTLIESQNOEKNOEQLLELDKMASLWNPFDITNNLMY 672

RESULT 3

envelope glycoprotein - human immunodeficiency virus type 1 (strain sc14.3)
T12016
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
Accession: T12016
R:McCuchan, F.E.; Sanders-Buell, E.; Salminen, M.O.; Carr, J.K.; Sheppard, W.H.
AIDS Res. Hum. Retroviruses 14, 329-337, 1998
A:Title: Diversity of the human immunodeficiency virus type 1 envelope glycoprotein in S
A:Reference number: Z17379; PMID:98178716; PMID:9519884
A:Accession: T12016
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-852 <MCC>
A:Cross-references: EMBL:U90934; NID:g2351783; PIDN:AAC59271.1; PID:g2351784
C:Genetic:
A:Gene: env
C:Superfamily: type E retrovirus env polypeptide

Query Match 86.9%; Score 3007.5; DB 2; Length 852;
Best Local Similarity 87.2%; Pred. No. 5e-209;
Matches 564; Conservative 31; Mismatches 47; Indels 5; Gaps 3;

Qy 2 EKLMTVYVYGVPMKEATTLTFCASDAKAYDEVHNVMAHACVPTDPNPOEVLNVTE 61
Db 31 EKLMTVYVYGVPMKEATTLTFCASDAKAYDEVHNVMAHACVPTDPNPOEVLNVTE 90
Qy 62 HENMKNMVEQNOEDIIISLDOSLKPCKVLTPLCVTLNCDVNAVNTTNDSEGMERGE 119
Db 91 HENMKNMVEQNOEDIIISLDOSLKPCKVLTPLCVTLNCDVNAVNTTNDSEGMERGE 150
Qy 120 GEIKCSFNITTSIRDEVOKEYALFKLDVVPIDNNNTSYRLISCDTSVITQACPKISFE 179
Db 151 GEIKCSFNITTSIRDEVOKEYALFKLDVVPIDNNNTSYRLISCDTSVITQACPKISFE 210
Qy 180 PIHYCAPAGFALIKCNDKTFNGKPCCKNVSTVQCTHGIRPVSTOLLNGLAEDEVVIR 239
Db 211 PIHYCAPAGFALIKCNDKTFNGKPCCKNVSTVQCTHGIRPVSTOLLNGLAEDEVVIR 270
Qy 240 IRSDNFTNNAKTIIVOLKESVEINCTRPNNNTKSHIIGRAFYTTGIIIGDIRAHGN 299
Db 271 IRSDNFTNNAKTIIVOLKESVEINCTRPNNNTKSHIIGRAFYTTGIIIGDIRAHGN 330
Qy 300 ISRAKNDTLKOIVILRBOFEKTIIVFNHSSGDEPEIWNHSGGEFFYCNSITGLLFT 359
Db 331 ISRAKNDTLKOIVILRBOFEKTIIVFNHSSGDEPEIWNHSGGEFFYCNSITGLLFT 390
Qy 360 TW--NNNTGSGNNTGNTIILPCRIKOIINMOEYKAMVAPIRGOIRCSNITGLLFT 416
Db 391 TW--NNNTGSGNNTGNTIILPCRIKOIINMOEYKAMVAPIRGOIRCSNITGLLFT 450
Qy 417 TRDGINENGTETFRPGGDMRDNRSELYKYKVKIEPLGVAPTKARRVOREKAVG 476
Db 451 TRDGINENGTETFRPGGDMRDNRSELYKYKVKIEPLGVAPTKARRVOREKAVG 510
Qy 477 IGAVFLGFLGAAGSTMGAASMTLTVQARLLLSGIYVOQNNLLRAIEAQORMLQLTWNGIK 536
Db 511 IGAVFLGFLGAAGSTMGAASMTLTVQARLLLSGIYVOQNNLLRAIEAQORMLQLTWNGIK 570
Qy 537 QLOARVLAVERYLDGDOQLIGIWCSSGKLICTTAVPWNASWSNKSIDRIIMNMNTWMEERE 596
Db 571 QLOARVLAVERYLDGDOQLIGIWCSSGKLICTTAVPWNASWSNKSIDRIIMNMNTWMEERE 630
Qy 597 IDNYTSEIYTLIESQNOEKNOEQLLELDKMASLWNPFDITNNLMY 643
Db 631 IDNYTSEIYTLIESQNOEKNOEQLLELDKMASLWNPFDITNNLMY 677

RESULT 4

env polyprotein precursor - human immunodeficiency virus type 1 (strain YU-2)

N1:Alternate names: coat polyprotein

N2:Contains: coat protein gp120; coat protein gp41

C1:Species: human immunodeficiency virus type 1, HIV-1

A1:Note: host Homo sapiens (man)

C1:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Sep-1994

C1:Accession: H44001

R1:Li, Y.; Hui, H.; Burgess, C.J.; Price, R.W.; Sharp, P.M.; Hahn, B.H.; Shaw, G.M.

J. Virol. 66, 6587-6600, 1992

A1:Title: Complete nucleotide sequence, genome organization, and biological properties of

A1:Reference number: A44001; MUID:93021387; PMID:1404605

A1:Accession: H44001

A1:Molecule type: DNA

A1:Residues: 1-843 <L1>

A1:Cross-References: GB:M93258

C1:Genetics:

C1:Gene: env

C1:Superfamily: type E retrovirus env polyprotein

C1:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein

F1:1-29/Domain: signal sequence #status predicted <SIG>

F1:19-35/Region: hydrophobic

F1:30-489/Product: coat protein gp120 #status predicted <GP1>

F1:490-843/Product: coat protein gp41 #status predicted <GP2>

F1:673-689/Region: hydrophobic

F1:738-755/Domain: transmembrane #status predicted <TM>

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

F1:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

Db 565 ARVLAVERYLDRDQQLGIWGSGLICTTAVPNAWSNKSIDRIINNTMMWEEREIDN 624

Qy 600 YTSSEYTLIESONOEKNEOELELDKWAALNMPDITNNLWY 643

Db 625 YTHIYSLIEOSONOEKNEOELELDKWAALNMPDITNNLWY 668

RESULT 5

env polyprotein precursor - human immunodeficiency virus type 1 (isolate LAV-1a)

N1:Alternate names: coat polyprotein

C1:Species: human immunodeficiency virus type 1, HIV-1

A1:Note: host Homo sapiens (man)

C1:Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 16-Jul-1999

C1:Accession: A03975

R1:Main-Hobson, S.; Sonigo, P.; Danos, O.; Cole, S.; Alizon, M.

Cell 40, 9-17, 1985

A1:Title: Nucleotide sequence of the AIDS virus, LAV.

A1:Reference number: A0866; MUID:85099333; PMID:2981635

A1:Accession: A03975

A1:Molecule type: DNA

A1:Residues: 1-861 <MAI>

A1:Cross-References: GB:K02013; NID:G326417; PID:AB59751.1; PID:G326424

C1:Genetics:

C1:Gene: env

C1:Superfamily: type E retrovirus env polyprotein

C1:Keywords: AIDS; Capaid protein; coat protein; glycoprotein; immunodeficiency; polyprot.

F1:1-30/Domain: signal sequence #status predicted <SIG>

F1:31-516/Product: exterior membrane glycoprotein #status predicted <EXT>

F1:517-861/Product: transmembrane glycoprotein #status predicted <TM>

F1:88,136,141,146,161,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411

F1:88,136,141,146,161,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411

F1:88,136,141,146,161,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411

F1:88,136,141,146,161,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411

F1:88,136,141,146,161,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411

F1:88,136,141,146,161,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411

F1:88,136,141,146,161,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411

F1:88,136,141,146,161,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411

F1:88,136,141,146,161,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411

F1:88,136,141,146,161,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411

F1:88,136,141,146,161,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411

F1:88,136,141,146,161,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411

F1:88,136,141,146,161,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411

F1:88,136,141,146,161,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411

F1:88,136,141,146,161,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411

F1:88,136,141,146,161,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411

F1:88,136,141,146,161,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411

F1:88,136,141,146,161,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411

F1:88,136,141,146,161,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411

F1:88,136,141,146,161,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411

F1:88,136,141,146,161,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411

F1:88,136,141,146,161,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411

F1:88,136,141,146,161,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411

F1:88,136,141,146,161,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411

F1:88,136,141,146,161,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411

F1:88,136,141,146,161,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411

F1:88,136,141,146,161,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411

F1:88,136,141,146,161,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411

F1:88,136,141,146,161,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411

F1:88,136,141,146,161,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411

F1:88,136,141,146,161,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411

F1:88,136,141,146,161,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411

F1:88,136,141,146,161,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411

F1:88,136,141,146,161,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411

QY	597	IDNTSEYTLIEESQNOQEKOELELDKMSLNMWEDITWLMY	643
Db	631	IDNTSEYTLIEESQNOQEKOELELDKMSLNMWEDITWLMY	677

RESULT 7

env polyprotein precursor - human immunodeficiency virus type 1 (isolate HTLV-III, BH10)

N:Alternate names: coat polyprotein

C:Species: human immunodeficiency virus type 1, HIV-1

A:Note: host Homo sapiens (man)

C:Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 16-Jul-1999

C:Accession: A03973

R:Author: J.A.; Hasseltine, W.; Patarca, R.; Liyak, K.J.; Starcich, B.; Josephs, S.F.; Doran, R.; Rabner, J.A.; Papas, T.S.; Graybe, J.; Chang, N.T.; Gallo, R.C.; Wong-Staal, F.

Feature 313, 277-284, 1985

A:Title: Complete nucleotide sequence of the AIDS virus, HTLV-III.

A:Reference number: A93353; M0ID:8511123; PMID:2578615

A:Accession: A03973

A:Molecule type: DNA

A:Residues: 1-856 <RAT>

C:Genetics:

A:Gene: env

C:Superfamily: type E retrovirus env polyprotein

C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprot.

F:1-30/Domains: signal sequence #status predicted <SIG>

F:31-511/Product: exterior membrane glycoprotein #status predicted <TM>

F:512-856/Product: transmembrane glycoprotein #status predicted <TM>

F:88,136,141,156,160,186,197,230,234,241,262,276,289,295,301,332,339,356,386,392,397,406

F:611,616,625,637,674,750,816/Binding site: carbohydrate (Asn) (covalent) #status predic

Query Match 86.0%; Score 2978; DB 1; Length 856;

Best Local Similarity 86.5%; Pred. No. 6 8e-207; Indels 10; Gaps 6;

Matches 563; Conservative 35; Mismatches 43;

QY	2	EKLMTVTVYVGVPMWEKATTTLPFCASDAKAYDEVHNVNMAHACVPTDPNPGVEVLENYTE	61
Db	32	EKLMTVTVYVGVPMWEKATTTLPFCASDAKAYDEVHNVNMAHACVPTDPNPGVEVLENYTE <td>91</td>	91

QY	62	HNMMKNMNVBEWQEDIIISLMQSLKPCVCLTFLCYTLNCKDNVATITNDSEG--THER	119
Db	92	NNMMKNMNVBEWQEDIIISLMQSLKPCVCLTFLCYTLNCKDNVATITNDSEG--THER <td>151</td>	151

QY	120	GEIKNCSFVITTSIRDEVOKEVALPFKLDVPPIDNNNTSYRLISCDTSVITQACPISFE	179
Db	152 <th>GEIKNCSFVITTSIRDEVOKEVALPFKLDVPPIDNNNTSYRLISCDTSVITQACPISFE</th> <th>211</th>	GEIKNCSFVITTSIRDEVOKEVALPFKLDVPPIDNNNTSYRLISCDTSVITQACPISFE	211

QY	180	PIPIHYCAPAGFAILKCNCKDTFNGKPCCKNVSTVOCTHGIRPVVSTOLLNGSLAEEBV	239
Db	212 <th>PIPIHYCAPAGFAILKCNCKDTFNGKPCCKNVSTVOCTHGIRPVVSTOLLNGSLAEEBV</th> <th>271</th>	PIPIHYCAPAGFAILKCNCKDTFNGKPCCKNVSTVOCTHGIRPVVSTOLLNGSLAEEBV	271

QY	240	IRSDNFTNNAKTIIVOLKKSVEINCTRPNNNTKSIHI--GPGAFPTTGIIIDIRAH	297
Db	272 <th>IRSAFTDPAKTIIVOLKKSVEINCTRPNNNTKSIHI--GPGAFPTTGIIIDIRAH</th> <th>330</th>	IRSAFTDPAKTIIVOLKKSVEINCTRPNNNTKSIHI--GPGAFPTTGIIIDIRAH	330

QY	298	CNISRAKNNDTLKQIYIKLREOF--ENKTIIVFNSSGGDPETIVMSFNCEGEFFYCNSTOL	356
Db	331 <th>CNISRAKNNDTLKQIYIKLREOF--ENKTIIVFNSSGGDPETIVMSFNCEGEFFYCNSTOL</th> <th>390</th>	CNISRAKNNDTLKQIYIKLREOF--ENKTIIVFNSSGGDPETIVMSFNCEGEFFYCNSTOL	390

QY	357	FNSTWNN--NTGSGNTEG--NTITLPCRIKQIINMVEGKAMVAPPINQIRCSSNIT	412
Db	391 <th>FNSTWNN--NTGSGNTEG--NTITLPCRIKQIINMVEGKAMVAPPINQIRCSSNIT</th> <th>450</th>	FNSTWNN--NTGSGNTEG--NTITLPCRIKQIINMVEGKAMVAPPINQIRCSSNIT	450

QY	413	GILLTRDGINENGTETFRPGGDMRDNRSEYKYVVKIETPLGVAPTKKRRVOREK	472
Db	451 <th>GILLTRDGINENGTETFRPGGDMRDNRSEYKYVVKIETPLGVAPTKKRRVOREK</th> <th>510</th>	GILLTRDGINENGTETFRPGGDMRDNRSEYKYVVKIETPLGVAPTKKRRVOREK	510

QY	473	RAVVGAVLGLFGAAGSTMGAASMTLVQARLLLSGIVOOQNNLLAIEAQOORLQTLV	532
Db	511 <th>RAVVGAVLGLFGAAGSTMGAASMTLVQARLLLSGIVOOQNNLLAIEAQOORLQTLV</th> <th>570</th>	RAVVGAVLGLFGAAGSTMGAASMTLVQARLLLSGIVOOQNNLLAIEAQOORLQTLV	570

QY 533 MGICLOARVLAVERYLDQDLGIGWSSGKLICTTAAPNPNASWNSKSLDRIMNNMTWME 592
 DB 571 MGICLOARVLAVERYLDQDLGIGWSSGKLICTTAAPNPNASWNSKSLDRIMNNMTWME 630
 QY 593 WEREIDNTSEIYLTIEESQNOEKNEOELLELDKMAISLMMNPDITNNLMY 643
 DB 631 WDREINNTSLIHSIEESQNOEKNEOELLELDKMAISLMMNPDITNNLMY 681

RESULT 8
 env polyprotein precursor - human immunodeficiency virus type 1 (isolate LV)

N.Alternate names: coat polyprotein
 C.Species: human immunodeficiency virus type 1, HIV-1
 A.Note: Host Homo sapiens (man)
 C.Date: 17-May-1985 #sequence_revision 17-May-1985 #ext_change 16-Jul-1999
 C.Accession: A03974
 R.Muehling, M.A.; Smith, D.H.; Cabradilla, C.D.; Benton, C.V.; Lasky, L.A.; Capon, D.J.
 Nature 313, 450-458, 1985
 A.Title: Nucleic acid structure and expression of the human AIDS/lymphadenopathy retrovirus
 A.Reference number: A93355; MUID:85111157; PMID:2982104
 A.Accession: A03974
 A.Molecule type: DNA
 A.Residues: 1-856 <MUB>
 A.Cross-references: GB:K02083; NID:9555008; PIDN:AAB59873.1; PID:9328559
 C.Gene: env
 C.Superfamily: type E retrovirus env polyprotein
 C.Keywords: Aids; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprot

F.1-30/Domain: -signal sequence #status predicted <Sig>
 F.31-51/Product: exterior membrane glycoprotein #status predicted <TM>
 F.512-856/Product: transmembrane glycoprotein #status predicted <TM>
 F.88,136,143,156,160,186,197,230,234,241,282,276,289,295,303,332,339,356,386,392,397,406
 F.611,616,625,637,674,750,816/Binding site: carbohydrate (Asn) (covalent) #status predic

Query Match 85.9%; Score 2974; DB 1; Length 856;
 Best Local Similarity 86.3%; Pred. No. 1.3e-206;
 Matches 562; Conservative 34; Mismatches 45; Indels 10; Gaps 6;

QY 2 EKLWTVVYGVVWKEATTLTFCASDAKAYDTEVNNWATACVPTDPNPQSEVLENTVE 61
 DB 32 EKLWTVVYGVVWKEATTLTFCASDAKAYDTEVNNWATACVPTDPNPQSEVLENTVE 91
 QY 62 HNNMKNMNVQEMOEDIISLMDQSLKPCVKLTPLCVTLNCKDVNATNTTNDSEG--TWER 119
 DB 92 HNNMKNMNVQEMOEDIISLMDQSLKPCVKLTPLCVTLNCKDVNATNTTNDSEG--TWER 151
 QY 120 GEIKNCSFNITTSIRDEVOKEVALFYKLDVVPIDNNNTSYRLISCDTSVITTOACPKISFE 179
 DB 152 GEIKNCSFNITTSIRDEVOKEVALFYKLDVVPIDNNNTSYRLISCDTSVITTOACPKISFE 211
 QY 180 PPIHYCAPAGFALLKCNKDTFNGKPCKNVSTVOCTHGIRPVVSTOLLNGLAEEV 239
 DB 212 PPIHYCAPAGFALLKCNKDTFNGKPCKNVSTVOCTHGIRPVVSTOLLNGLAEEV 271
 QY 240 IRSDFTNNAKTIIVQLKESVEINCTRPNNNTRKSIH--GGRAFYTTGELIGDIRQAH 297
 DB 272 IRSDFTNNAKTIIVQLKESVEINCTRPNNNTRKSIH--GGRAFYTTGELIGDIRQAH 330
 QY 298 CNISRAKNNDTLKQIVIKLREOF-ENKTIIVNHSGGDEPEIVMHSFNGEGEFYCNSTOL 356
 DB 331 CNISRAKNNDTLKQIVIKLREOF-ENKTIIVNHSGGDEPEIVMHSFNGEGEFYCNSTOL 390
 QY 357 FNSTNN--NTEGSNNTEG-NTITLPCRIOIIMMOEYGVKAMAPPIRQIRCSNIT 412
 DB 391 FNSTNN--NTEGSNNTEG-NTITLPCRIOIIMMOEYGVKAMAPPIRQIRCSNIT 450
 QY 413 GLLTRDGGINENGTEIFRRPGGDMRDNRSEFYKVKVKEPLGLVAATPKAKRRVVOREK 472
 DB 451 GLLTRDGGINENGTEIFRRPGGDMRDNRSEFYKVKVKEPLGLVAATPKAKRRVVOREK 510
 QY 473 RAVGIGAVFLGFLGAAGSTMGAAASMTLTVQARLLLSGIYVQOONNLRAIEAQRMQLTV 532

DB 511 RAVGIGAVFLGFLGAAGSTMGAAASMTLTVQARLLLSGIYVQOONNLRAIEAQRMQLTV 570
 QY 533 MGICLOARVLAVERYLDQDLGIGWSSGKLICTTAAPNPNASWNSKSLDRIMNNMTWME 592
 DB 571 MGICLOARVLAVERYLDQDLGIGWSSGKLICTTAAPNPNASWNSKSLDRIMNNMTWME 630
 QY 593 WEREIDNTSEIYLTIEESQNOEKNEOELLELDKMAISLMMNPDITNNLMY 643
 DB 631 WDREINNTSLIHSIEESQNOEKNEOELLELDKMAISLMMNPDITNNLMY 681

RESULT 9

env protein - human immunodeficiency virus type 1
 C.Species: human immunodeficiency virus type 1, HIV-1
 C.Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #ext_change 31-Oct-1997
 C.Accession: S13288
 R.O'Brien, W.A.; Koyanagi, Y.; Namazie, A.; Zhao, J.Q.; Diagne, A.; Idler, K.; Zack, J.A.
 Nature 348, 69-73, 1990
 A.Title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120
 A.Reference number: S13288; MUID:91043044; PMID:2172833
 A.Accession: S13288
 A.Status: preliminary
 A.Molecule type: DNA
 A.Residues: 1-854 <OBR>
 C.Superfamily: type E retrovirus env polyprotein

Query Match 85.8%; Score 2972; DB 2; Length 854;
 Best Local Similarity 86.6%; Pred. No. 1.8e-206;
 Matches 564; Conservative 30; Mismatches 45; Indels 12; Gaps 7;

QY 2 EKLWTVVYGVVWKEATTLTFCASDAKAYDTEVNNWATACVPTDPNPQSEVLENTVE 61
 DB 32 EKLWTVVYGVVWKEATTLTFCASDAKAYDTEVNNWATACVPTDPNPQSEVLENTVE 91
 QY 62 HNNMKNMNVQEMOEDIISLMDQSLKPCVKLTPLCVTLNCKDVNATNTTNDSEG--TWER 119
 DB 92 HNNMKNMNVQEMOEDIISLMDQSLKPCVKLTPLCVTLNCKDVNATNTTNDSEG--TWER 151
 QY 120 GEIKNCSFNITTSIRDEVOKEVALFYKLDVVPIDNNNTSYRLISCDTSVITTOACPKISFE 179
 DB 152 GEIKNCSFNITTSIRDEVOKEVALFYKLDVVPIDNNNTSYRLISCDTSVITTOACPKISFE 209
 QY 180 PPIHYCAPAGFALLKCNKDTFNGKPCKNVSTVOCTHGIRPVVSTOLLNGLAEEV 239
 DB 210 PPIHYCAPAGFALLKCNKDTFNGKPCKNVSTVOCTHGIRPVVSTOLLNGLAEEV 269
 QY 240 IRSDFTNNAKTIIVQLKESVEINCTRPNNNTRKSIH--GGRAFYTTGELIGDIRQAH 297
 DB 270 IRSDFTNNAKTIIVQLKESVEINCTRPNNNTRKSIH--GGRAFYTTGELIGDIRQAH 328
 QY 298 CNISRAKNNDTLKQIVIKLREOF-ENKTIIVNHSGGDEPEIVMHSFNGEGEFYCNSTOL 356
 DB 329 CNISRAKNNDTLKQIVIKLREOF-ENKTIIVNHSGGDEPEIVMHSFNGEGEFYCNSTOL 388
 QY 357 FNSTNN--NTEGSNNTEG-NTITLPCRIOIIMMOEYGVKAMAPPIRQIRCSNIT 412
 DB 389 FNSTNN--NTEGSNNTEG-NTITLPCRIOIIMMOEYGVKAMAPPIRQIRCSNIT 448
 QY 413 GLLTRDGGINENGTEIFRRPGGDMRDNRSEFYKVKVKEPLGLVAATPKAKRRVVOREK 472
 DB 449 GLLTRDGGINENGTEIFRRPGGDMRDNRSEFYKVKVKEPLGLVAATPKAKRRVVOREK 508
 QY 473 RAVGIGAVFLGFLGAAGSTMGAAASMTLTVQARLLLSGIYVQOONNLRAIEAQRMQLTV 532
 DB 509 RAVGIGAVFLGFLGAAGSTMGCTSWTLTVQARLLLSGIYVQOONNLRAIEAQRMQLTV 568
 QY 533 MGICLOARVLAVERYLDQDLGIGWSSGKLICTTAAPNPNASWNSKSLDRIMNNMTWME 592
 DB 569 MGICLOARVLAVERYLDQDLGIGWSSGKLICTTAAPNPNASWNSKSLDRIMNNMTWME 628
 QY 593 WEREIDNTSEIYLTIEESQNOEKNEOELLELDKMAISLMMNPDITNNLMY 643

Db 629 WDREINNTSILSHLIESQNOEKNEBELLELDKMSLMMFNTLWMT 679

RESULT 10
VCLJAJ2
env polyprotein precursor - human immunodeficiency virus type 1 (isolate ARV-2)

N:Alternate names: coat polyprotein
C:Species: human immunodeficiency virus type 1, HIV-1
A:Note: host Homo sapiens (man)
C:Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 16-Jul-1999
C:Accession: A03976
R:Sanchez-Pescador, R.; Power, M.D.; Barr, P.J.; Steimer, K.S.; Stempien, M.M.; Brown-St
A:Title: Nucleotide sequence and expression of an AIDS-associated retrovirus (ARV-2).
A:Reference number: A04003; MUID:85090453; PMID:2578227
A:Accession: A03976
A:Molecule type: DNA
A:Residues: 1-855
A:References: GB:X02007; NID:G328658; PID:AA859882.1; PID:G328666

A:Gene: env
C:Superfamily: type E retrovirus env polyprotein
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprot
F:1-30/Domain: signal sequence #status predicted <SIG>
F:31-509/Product: exterior membrane glycoprotein #status predicted <EXT>
F:510-855/Product: transmembrane glycoprotein #status predicted <TM>
F:87,129,140,158,184,190,200,244,265,282,298,304,334,351,358,364,388,394,400,408,445,458
F:610,624,636,815/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 85.8%; Score 2972; DB 1; Length 855;
Best Local Similarity 86.2%; Pred. No. 1.8e-206;
Matches 563; Conservative 32; Mismatches 44; Indels 14; Gaps 8;

Db 2 EKLMTVTVYGVVWKEATTTLLFCASDAKAYDEVHNMWATHACVPTDNPQEVLENTTE 61
Db 31 EKLMTVTVYGVVWKEATTTLLFCASDAKAYDEVHNMWATHACVPTDNPQEVLENTTE 90
Qy 62 HFNMKNMNMWQOMEDIIISLMDOSLKPVCVLTPLCTVLTNCKDVANTNTNDSGME-RG 120
Db 91 HFNMKNMNMWQOMEDIIISLMDOSLKPVCVLTPLCTVLTNCKDVANTNTNDSGME-RG 150
Qy 121 EIKNCSFNITISIRDEVQKALFYKLDVPEIDN-----NNTSYRLISCDTSVITQACPK 175
Db 151 EIKNCSFNITISIRDEVQKALFYKLDVPEIDN-----NNTSYRLISCDTSVITQACPK 210
Qy 176 ISFEPDPIHYCAPAGFALIKNDKTFNGKPGCKNVSTVOCTHGIRPVVSTOLLNGSL 235
Db 211 VSFEPDPIHYCAPAGFALIKNDKTFNGKPGCKNVSTVOCTHGIRPVVSTOLLNGSL 270
Qy 236 EEVVIRSNFTNNAKTIIIVOLKESVEINCTRPNNTRKSIHIGPRAFTTGEIIGDI 295
Db 271 EEVVIRSNFTNNAKTIIIVOLKESVEINCTRPNNTRKSIHIGPRAFTTGEIIGDI 330
Qy 296 AHCNISRAKMDTLKQIYIKLEQEF-ENKTIIVFNHSGGDEIIVHNSFCGEGFFYCNST 354
Db 331 AHCNISRAKMDTLKQIYIKLEQEF-ENKTIIVFNHSGGDEIIVHNSFCGEGFFYCNST 390
Qy 355 QLFNSTWNN--TESSNNT-EGNTITLPCRIOIINMGOEVKANYAPPIRQIQCSS 412
Db 391 QLFNSTWNN--TESSNNT-EGNTITLPCRIOIINMGOEVKANYAPPIRQIQCSS 447
Qy 413 GLLLTLTRDGG--INSGTEIFPRGGGDMRDMNRSEFYKYKVKIEPLGVAFTCKRRV 471
Db 448 GLLLTLTRDGG--INSGTEIFPRGGGDMRDMNRSEFYKYKVKIEPLGVAFTCKRRV 507
Qy 472 KRAVGI-GAVELGFLGAAGSTGASMTLTVQARLLLSGIVQOQNNLLRAIEAOQRM 530
Db 508 KRAVGI-GAVELGFLGAAGSTGASMTLTVQARLLLSGIVQOQNNLLRAIEAOQRM 567
Qy 531 TWGIGIKOLQARVLAVERLYGDOQLLGWCGSGKLICTTAVPWNASWSKSLDRINWNTW 590
Db 568 TWGIGIKOLQARVLAVERLYGDOQLLGWCGSGKLICTTAVPWNASWSKSLDRINWNTW 627

Qy 591 MEMEREIDNTSITLIESQNOEKNEBELLELDKMSLMMFNTLWMT 643
Db 628 MEMEREIDNTSITLIESQNOEKNEBELLELDKMSLMMFNTLWMT 680

RESULT 11

VCLJMN
env polyprotein precursor - human immunodeficiency virus type 1 (isolate MN)

N:Alternate names: coat polyprotein
C:Species: human immunodeficiency virus type 1, HIV-1
A:Note: host Homo sapiens (man)
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Feb-1997
C:Accession: A28922
R:Gungo, C.; Guo, H.G.; Franchini, G.; Aldovini, A.; Collalti, E.; Farrell, K.; Wong-Stat
A:Title: Envelope sequences of two new United States HIV-1 isolates.
A:Reference number: A28922; MUID:88219542; PMID:3369091
A:Accession: A28922
A:Molecule type: DNA
A:Residues: 1-859 <CUR>

A:Gene: env
C:Superfamily: type E retrovirus env polyprotein
C:Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane protei
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-859/Product: env polyprotein #status predicted <EP>
F:87,129,135,140,141,146,161,165,191,202,246,261,294,300,336,343,359,365,389,395,401

Query Match 85.7%; Score 2967.5; DB 1; Length 859;
Best Local Similarity 85.7%; Pred. No. 3.9e-206;
Matches 562; Conservative 32; Mismatches 47; Indels 15; Gaps 6;

Db 2 EKLMTVTVYGVVWKEATTTLLFCASDAKAYDEVHNMWATHACVPTDNPQEVLENTTE 61
Db 31 EKLMTVTVYGVVWKEATTTLLFCASDAKAYDEVHNMWATHACVPTDNPQEVLENTTE 90
Qy 62 HFNMKNMNMWQOMEDIIISLMDOSLKPVCVLTPLCTVLTNCKDVANTNTNDSGME-RG 113
Db 91 HFNMKNMNMWQOMEDIIISLMDOSLKPVCVLTPLCTVLTNCKDVANTNTNDSGME-RG 150
Qy 114 EGMERGEIKNCSFNITISIRDEVQKALFYKLDVPEIDNNTSYRLISCDTSVITQAC 173
Db 151 EGMERGEIKNCSFNITISIRDEVQKALFYKLDVPEIDNNTSYRLISCDTSVITQAC 210
Qy 174 PKISFEPDPIHYCAPAGFALIKNDKTFNGKPGCKNVSTVOCTHGIRPVVSTOLLNGSL 233
Db 211 PKISFEPDPIHYCAPAGFALIKNDKTFNGKPGCKNVSTVOCTHGIRPVVSTOLLNGSL 270
Qy 234 AEEVIRSNFTNNAKTIIIVOLKESVEINCTRPNNTRKSIHIGPRAFTTGEIIGDI 293
Db 271 AEEVIRSNFTNNAKTIIIVOLKESVEINCTRPNNTRKSIHIGPRAFTTGEIIGDI 330
Qy 294 ROAHNISRAKMDTLKQIYIKLEQEF-ENKTIIVFNHSGGDEIIVHNSFCGEGFFYCNST 353
Db 331 ROAHNISRAKMDTLKQIYIKLEQEF-ENKTIIVFNHSGGDEIIVHNSFCGEGFFYCNST 390
Qy 354 TOLFNSTWNN--TESSNNT-EGNTITLPCRIOIINMGOEVKANYAPPIRQIQCSS 409
Db 391 TOLFNSTWNN--TESSNNT-EGNTITLPCRIOIINMGOEVKANYAPPIRQIQCSS 450
Qy 410 NITGLLTLTRDGG--INSGTEIFPRGGGDMRDMNRSEFYKYKVKIEPLGVAFTCKRRV 467
Db 451 NITGLLTLTRDGG--INSGTEIFPRGGGDMRDMNRSEFYKYKVKIEPLGVAFTCKRRV 510
Qy 468 VOREKRAVGI-GAVELGFLGAAGSTGASMTLTVQARLLLSGIVQOQNNLLRAIEAOQRM 527
Db 511 VOREKRAVGI-GAVELGFLGAAGSTGASMTLTVQARLLLSGIVQOQNNLLRAIEAOQRM 569
Qy 528 LQLTWVGIKOLQARVLAVERLYGDOQLLGWCGSGKLICTTAVPWNASWSKSLDRINWNTW 587
Db 570 LQLTWVGIKOLQARVLAVERLYGDOQLLGWCGSGKLICTTAVPWNASWSKSLDRINWNTW 629

QY 568 MTWMEEREIDNYTSEIYTLIEESQNOQEKNEOELLELDKWSLMMNPDITNMLWY 643
 DB 630 MTWMEEREIDNYTSEIYTLIEESQNOQEKNEOELLELDKWSLMMNPDITNMLWY 685

RESULT 12 VCLJ3W

env polyprotein precursor - human immunodeficiency virus type 1 (isolate SC)
 N:Alternate names: coat protein polyprotein
 C:Species: human immunodeficiency virus type 1, HIV-1
 A:Note: host Homo sapiens (man)

C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Feb-1997
 C:Accession: B28922
 R:Guigo, C.; Guo, H.-G.; Franchini, G.; Aldovini, A.; Collalti, E.; Farrell, K.; Wong-Ste
 Virology 164, 531-536, 1988
 A:Title: Envelope sequences of two new United States HIV-1 isolates.
 A:Reference number: A28922; MUID:88219542; PMID:3369091
 A:Accession: B28922
 A:Molecule type: DNA
 A:Residues: 1-861 <GUR>

A:Gene: env

C:Superfamily: type E retrovirus env polyprotein
 C:Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane prote
 F:1-39/Domain: signal sequence #status predicted <SIG>
 F:30-861/Product: env polyprotein #status predicted <EP>
 F:87,129,135,140,143,159,163,187,198,234,241,262,276,295,301,302,333,340,356,362,386,396

Query Match 85.7%; Score 2967; DB 1; Length 861;
 Best Local Similarity 85.6%; Pred. No. 4.2e-206;
 Matches 565; Conservative 34; Mismatches 39; Indels 22; Gaps 9;

QY 2 EKLMTVTVYGVVPMKEATTTTLFCASDAKAYDTEVHNWATHACVPTDPOEVLENT 61
 DB 31 EQLMTVTVYGVVPMKEATTTTLFCASDAKAYDTEVHNWATHACVPTDPOEVLENT 90
 QY 62 HENMMKNNMVEQMOEDIIISLMDQSLKPCVLTPLCTVLTNCQDV-----NATNTTNDSEG 115
 DB 91 HENMMKNNMVEQMOEDIIISLMDQSLKPCVLTPLCTVLTNCQDV-----NATNTTNDSEG 150
 QY 116 TMEBEIKKCSFNITTSIRDEQKEYALFYKLDVVPIDNNNTSYRLISCDTSVITQACP 175
 DB 151 KMEGEMNCSFNITTSIRDEQKEYALFYKLDVVPIDNNNTSYRLISCDTSVITQACP 208
 QY 176 ISFEPIPIHYCAPAGFALIKNDKTFNGKPCQKAVSTVOCTHGIRPVYSTOLLNGSLAE 235
 DB 209 VSFEPIPIHYCAPAGFALIKNDKTFNGKPCQKAVSTVOCTHGIRPVYSTOLLNGSLAE 267
 QY 236 EEVVIRSDNFTNNAKTIIVQLKESVEINCTRPNNTRKSIH--GPGRAFYTTGEBIIGDI 293
 DB 268 EEVVIRSENFDTNNAKTIIVQLKESVEINCTRPNNTRKSIH--GPGRAFYTTGEBIIGDI 327
 QY 294 QAHNCNISRAKXNDTLKQIVIKLRQFENKTIIVFNHSSGDDPEIYHNSFCGEFFYCNST 353
 DB 328 QAHNCNISRAKXNDTLKQIVIKLRQFENKTIIVFNHSSGDDPEIYHNSFCGEFFYCNST 387
 QY 354 TQLFNSTNNNTGEGS---NTEGNT--TTLPCRIOIIMMOGVGKAMAPRIRGDIRCSS 409
 DB 388 TQLFNSTNNNTGEGS---NTEGNT--TTLPCRIOIIMMOGVGKAMAPRIRGDIRCSS 446
 QY 410 NITGLLTRDGINENG---TEIFRPGGDMRDMRSEFYKYKVKIEPLGVAFTCK 464
 DB 447 NITGLLTRDGINENG---TEIFRPGGDMRDMRSEFYKYKVKIEPLGVAFTCK 506
 QY 465 RRVVREKRAVG--IGAVPLGFLGAASTWGAASMTLTVQARLLLSGIYQOONNLRAIEA 523
 DB 507 RRVVREKRAVG--IGAVPLGFLGAASTWGAASMTLTVQARLLLSGIYQOONNLRAIEA 566
 QY 524 QORMLQITWGIKQIQAARLAVERYLGDQQLLGIGWCSGKLCCTAVPNAASNSKSLDR 583
 DB 567 QORMLQITWGIKQIQAARLAVERYLGDQQLLGIGWCSGKLCCTAVPNAASNSKSLDR 626
 QY 584 IWNMTWMEEREIDNYTSEIYTLIEESQNOQEKNEOELLELDKWSLMMNPDITNMLWY 643

DB 627 IWNMTWMEEREIDNYTSEIYTLIEESQNOQEKNEOELLELDKWSLMMNPDITNMLWY 686

RESULT 13 VCLJ3W

env polyprotein precursor - human immunodeficiency virus type 1 (isolate WMJ1)
 N:Contains: coat protein gp120; coat protein gp41
 C:Species: human immunodeficiency virus type 1, HIV-1
 C:Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 07-Nov-1997

C:Accession: A24774
 R:Starcich, B.R.; Hahn, B.H.; Shaw, G.M.; McNeely, P.D.; Modrow, S.; Wolf, H.; Parks, E
 Cell 45, 637-648, 1986
 A:Title: Identification and characterization of conserved and variable regions in the en
 A:Reference number: A24774; MUID:86218077; PMID:2423250
 A:Accession: A24774
 A:Molecule type: DNA
 A:Residues: 1-856 <STA>
 A:Cross-references: GB:K03455; GB:M38432; NID:g1906382
 C:Genetics:

A:Gene: env

C:Superfamily: type E retrovirus env polyprotein
 C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein
 F:1-39/Domain: signal sequence #status predicted <SIG>
 F:30-501/Product: coat protein gp120 #status predicted <GP1>
 F:502-847/Product: coat protein gp41 #status predicted <GP2>
 F:87,134,140,151,155,183,197,234,241,262,276,289,295,331,338,354,360,390,394,404,447,459

Query Match 85.5%; Score 2958.5; DB 1; Length 856;
 Best Local Similarity 86.1%; Pred. No. 1.7e-205;
 Matches 564; Conservative 30; Mismatches 46; Indels 15; Gaps 7;

QY 1 VEKLMTVTVYGVVPMKEATTTTLFCASDAKAYDTEVHNWATHACVPTDPOEVLENT 60
 DB 30 VEKLMTVTVYGVVPMKEATTTTLFCASDAKAYDTEVHNWATHACVPTDPOEVLENT 89
 QY 61 EHFNMKNNMVEQMOEDIIISLMDQSLKPCVLTPLCTVLTNCQDVATNTTNDSEGTMERG 120
 DB 90 EHFNMKNNMVEQMOEDIIISLMDQSLKPCVLTPLCTVLTNCQDVATNTTNDSEGTMERG 147
 QY 121 EIKNCSFNITTSIRDEQKEYALFYKLDVVPIDNNNTSYRLISCDTSVITQACP 175
 DB 148 EIKNCSFNITTSIRDEQKEYALFYKLDVVPIDNNNTSYRLISCDTSVITQACP 207
 QY 176 ISFEPIPIHYCAPAGFALIKNDKTFNGKPCQKAVSTVOCTHGIRPVYSTOLLNGSLAE 235
 DB 208 VSFEPIPIHYCAPAGFALIKNDKTFNGKPCQKAVSTVOCTHGIRPVYSTOLLNGSLAE 267
 QY 236 EEVVIRSDNFTNNAKTIIVQLKESVEINCTRPNNTR--RKSIIHGPAFYTTGEBIIGDI 294
 DB 268 EEVVIRSENFDTNNAKTIIVQLKESVEINCTRPNNTR--RKSIIHGPAFYTTGEBIIGDI 326
 QY 295 QAHNCNISRAKXNDTLKQIVIKLRQFENKTIIVFNHSSGDDPEIYHNSFCGEFFYCNST 354
 DB 327 QAHNCNISRAKXNDTLKQIVIKLRQFENKTIIVFNHSSGDDPEIYHNSFCGEFFYCNST 386
 QY 355 QLFNSTNNNTGEGS---NTEGNT--TTLPCRIOIIMMOGVGKAMAPRIRGDIRCSS 409
 DB 387 QLFNSTNNNTGEGS---NTEGNT--TTLPCRIOIIMMOGVGKAMAPRIRGDIRCSS 446
 QY 410 NITGLLTRDGINENG---TEIFRPGGDMRDMRSEFYKYKVKIEPLGVAFTCKGRVYQ 469
 DB 447 NITGLLTRDGINENG---TEIFRPGGDMRDMRSEFYKYKVKIEPLGVAFTCKGRVYQ 506
 QY 470 REKRAVG--IGAVPLGFLGAASTWGAASMTLTVQARLLLSGIYQOONNLRAIEAQMRL 528
 DB 507 REKRAVG--IGAVPLGFLGAASTWGAASMTLTVQARLLLSGIYQOONNLRAIEAQMRL 566
 QY 529 QITWGIKQIQAARLAVERYLGDQQLLGIGWCSGKLCCTAVPNAASNSKSLDR 588
 DB 567 QITWGIKQIQAARLAVERYLGDQQLLGIGWCSGKLCCTAVPNAASNSKSLDR 626
 QY 589 TMEWEEREIDNYTSEIYTLIEESQNOQEKNEOELLELDKWSLMMNPDITNMLWY 643

Db 627 TWMEWERIDNVTSLIVTLIESQNOQEKNOELLELDKMASLNMWFTITWLMY 681

RESULT 14

S33985

env polyprotein - human immunodeficiency virus type 1

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 26-Aug-1999

C:Accession: S33985

R:Carlani, F.

submitted to the EMBL Data Library, November 1991

A:Reference number: S33979

A:Accession: S33985

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-851 <CAR>

A:Cross-references: EMBL:Z11530; NID:G60192; PIDN:CAA77628.1; PID:G60199

C:Superfamily: type E retrovirus env polyprotein

Query Match 85.3%; Score 2953.5; DB 2; Length 851;

Best Local Similarity 86.3%; Pred. No. 3.9e-205;

Matches 559; Conservative 34; Mismatches 46; Indels 9; Gaps 6;

Qy 2 EKLWTVYGVVWPKKATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLEANTE 61
Db 32 EKLWTVYGVVWPKKATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLEANTE 91

Qy 62 HFNWKNMVEQMOEDIIISLMDOSLKPVCVLTPLCVTLNCKDVNATNTDSEGTMER 119
Db 92 HFNWKNMVEQMOEDIIISLMDOSLKPVCVLTPLCVTLNCKDVNATNTDSEGTMER 151

Qy 120 GEINCSFNITTSIRDEVOKEYALFYKLDVVPIDNNNTSYLLISCDTSVITQACPXSFE 179
Db 152 GEINCSFNITTSIRDEVOKEYALFYKLDVVPIDNNNTSYLLISCDTSVITQACPXSFE 211

Qy 180 PIPHYCAPAGFALIKCNDKTFNGKPCKNSTVQCTGIRPVVSTOLLNGLSLAESEV 239
Db 212 PIPHYCAPAGFALIKCNDKTFNGKPCKNSTVQCTGIRPVVSTOLLNGLSLAESEV 271

Qy 240 IRSNFTNNAKTIIVOLKESVEINCTRPNNNTKRSIH-GRGAFYTTGSHIGDIRQAH 297
Db 272 IRSNFTNNAKTIIVOLKESVEINCTRPNNNTKRSIH-GRGAFYTTGSHIGDIRQAH 330

Qy 298 CNISRAKNDTLKOIVIKRBOFENKTIIVNHSFGDPEIVMHSFNGEGFFYCNSTQL 356
Db 331 CNISRAKNDTLKOIVIKRBOFENKTIIVNHSFGDPEIVMHSFNGEGFFYCNSTQL 390

Qy 357 FNSTNNNTGSSNNTG-NTITLPCRIOIIMMOEVGKAMAPPIRGQIRCSSNITGL 415
Db 391 FNSTNNNTGSSNNTG-NTITLPCRIOIIMMOEVGKAMAPPIRGQIRCSSNITGL 448

Qy 416 LTRDGINENGTETFRPGGDMRDNRSEFYKYVVKIEPLGVAFTCKRRVVOREKAV 475
Db 449 LTRDGINENGTETFRPGGDMRDNRSEFYKYVVKIEPLGVAFTCKRRVVOREKAV 508

Qy 476 GIGAVFLGFLGAGSTWGAASMTLVQARLLLSGIVQOONNLRAIEAOQMLQUTWGI 535
Db 509 GIGAVFLGFLGAGSTWGAASMTLVQARLLLSGIVQOONNLRAIEAOQMLQUTWGI 568

Qy 536 KQOLARVLAVERYLQDQOLGIMWCGSKLICTTAVPNAASNSKSLDIRIMNMWTEWER 595
Db 569 KQOLARVLAVERYLQDQOLGIMWCGSKLICTTAVPNAASNSKSLDIRIMNMWTEWER 628

Qy 596 EIDNVTSEIYTLIESQNOQEKNOELLELDKMASLNMWFTITWLMY 643
Db 629 EIDNVTSEIYTLIESQNOQEKNOELLELDKMASLNMWFTITWLMY 676

RESULT 15

VCLIKX

env polyprotein precursor - human immunodeficiency virus type 1 (strain KB-1-gp32)

N:Alternate names: coat polyprotein

N:Contains: coat protein gp120; coat protein gp32

C:Species: human immunodeficiency virus type 1, HIV-1

A:Name: host Homo sapiens (man)

C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 01-Mar-1996

C:Accession: B42995

R:Shimizu, H.; Hasebe, F.; Tsuchie, H.; Morikawa, S.; Uehijima, H.; Kitamura, T.

Virology 189, 534-546, 1992

A:Title: Analysis of a human immunodeficiency virus type 1 isolate carrying a truncated

A:Reference number: A42995; MUID:92351552; PMID:1322587

A:Accession: B42995

A:Molecule type: mRNA

A:Residues: 1-729 <SHI>

A:Cross-references: GB:S41266; GB:D01206

C:Genetics:

C:Superfamily: type E retrovirus env polyprotein

C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein

F:1-689/Domain: extracellular #status predicted <EXT>

F:1-33/Domain: signal sequence #status predicted <SIG>

F:1-33/Region: hydrophobic #status predicted

F:34-517/Product: coat protein gp120 #status predicted <CP1>

F:514-517/Region: cleavage processing #status predicted <CP2>

F:518-729/Product: coat protein gp32 #status predicted <CP2>

F:518-534/Region: hydrophobic #status predicted

F:690-711/Domain: transmembrane #status predicted <TM1>

F:712-729/Domain: intracellular #status predicted <INT>

F:93.141,145,146,163,191,192,227,241,248,269,283,296,308,338,345,361,367,397,403,408,414,

Query Match 83.6%; Score 2895.5; DB 1; Length 729;

Best Local Similarity 82.8%; Pred. No. 4.9e-201;

Matches 539; Conservative 40; Mismatches 63; Indels 9; Gaps 2;

Qy 2 EKLWTVYGVVWPKKATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLEANTE 61
Db 37 EKLWTVYGVVWPKKATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLEANTE 96

Qy 62 HFNWKNMVEQMOEDIIISLMDOSLKPVCVLTPLCVTLNCKDVNATNTDSEGTMERGE 121
Db 97 HFNWKNMVEQMOEDIIISLMDOSLKPVCVLTPLCVTLNCKDVNATNTDSEGTMERGE 156

Qy 122 IKNCSEFNITTSIRDEVOKEYALFYKLDVVPIDNN-NTSYLLISCDTSVITQACPXS 177
Db 157 IKNCSEFNITTSIRDEVOKEYALFYKLDVVPIDNNNTSYLLISCDTSVITQACPXS 216

Qy 178 PEPPIHYCAPAGFALIKCNDKTFNGKPCKNSTVQCTGIRPVVSTOLLNGLSLAESE 237
Db 217 PEPPIHYCAPAGFALIKCNDKTFNGKPCKNSTVQCTGIRPVVSTOLLNGLSLAESE 276

Qy 238 VTRSDNFTNNAKTIIVOLKESVEINCTRPNNNTKRSIHIGPRAFYTTGSHIGDIRQAH 297
Db 277 VTRSDNFTNNAKTIIVOLKESVEINCTRPNNNTKRSIHIGPRAFYTTGSHIGDIRQAH 336

Qy 298 CNISRAKNDTLKOIVIKRBOFENKTIIVNHSFGDPEIVMHSFNGEGFFYCNSTQL 357
Db 337 CNISRAKNDTLKOIVIKRBOFENKTIIVNHSFGDPEIVMHSFNGEGFFYCNSTQL 396

Qy 358 NSTNNNTGSSNNTG-NTITLPCRIOIIMMOEVGKAMAPPIRGQIRCSSNITGL 417
Db 397 NSTNNNTGSSNNTG-NTITLPCRIOIIMMOEVGKAMAPPIRGQIRCSSNITGL 456

Qy 418 RDGINENGT-----EIFRPGGDMRDNRSEFYKYVVKIEPLGVAFTCKRRVVOREK 472
Db 457 RDGINENGT-----EIFRPGGDMRDNRSEFYKYVVKIEPLGVAFTCKRRVVOREK 516

Qy 473 RAVGIGAVFLGFLGAGSTWGAASMTLVQARLLLSGIVQOONNLRAIEAOQMLQUTV 532
Db 517 RAVGIGAVFLGFLGAGSTWGAASMTLVQARLLLSGIVQOONNLRAIEAOQMLQUTV 576

Qy 533 WGIKQOLARVLAVERYLQDQOLGIMWCGSKLICTTAVPNAASNSKSLDIRIMNMWTE 592
Db 577 WGIKQOLARVLAVERYLQDQOLGIMWCGSKLICTTAVPNAASNSKSLDIRIMNMWTE 636

Qy 593 WEREIDNVTSEIYTLIESQNOQEKNOELLELDKMASLNMWFTITWLMY 643

Db 637 WEREINNYTNLYNLIBESONQOEKNEODLALDKWDSILNWFSTKXLMY 687

Search completed: December 12, 2003, 12:33:31
Job time: 16.9697 secs

THIS PAGE BLANK (US.710)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: December 12, 2003, 12:25:38 ; Search time 10.7921 Seconds
(without alignments)
2801.881 Million cell updates/sec

Title: US-10-032-162-13

Perfect score: 3462

Sequence: 1 VEKLMVTYGVGPVWKEATL.....ELDKWASIMNFDITNLMY 643

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

1 number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3139.5	90.7	848	1 ENV_HV1JR	P20871 human immun
2	3041	87.8	847	1 ENV_HV1S1	P19550 human immun
3	3027.5	87.4	856	1 ENV_HV1SC	P05878 human immun
4	3000	86.7	843	1 ENV_HV1Y2	P35961 human immun
5	2997	86.6	867	1 ENV_HV1J3	P12489 human immun
6	2993.5	86.4	861	1 ENV_HV1BR	P03377 human immun
7	2992	86.4	852	1 ENV_HV1BN	P12488 human immun
8	2978	86.0	856	1 ENV_HV1BN	P03375 human immun
9	2974	85.9	856	1 ENV_HV1MN	P05877 human immun
10	2974	85.8	855	1 ENV_HV1PV	P03376 human immun
11	2972	85.8	855	1 ENV_HV1A2	P03378 human immun
12	2964	85.6	856	1 ENV_HV1H2	P04578 human immun
13	2958.5	85.5	856	1 ENV_HV1W1	P31872 human immun
14	2954.5	85.3	856	1 ENV_HV1H3	P04624 human immun
15	2954.5	85.2	851	1 ENV_HV1S3	P19549 human immun
16	2948.5	85.2	851	1 ENV_HV1B8	P04582 human immun
17	2946	85.1	856	1 ENV_HV1LW	O70626 human immun
18	2921	84.4	853	1 ENV_HV1MF	P19551 human immun
19	2917	84.3	855	1 ENV_HV1OY	P20888 human immun
20	2910	84.3	847	1 ENV_HV1RH	P04579 human immun
21	2895.5	83.6	861	1 ENV_HV1W2	P05880 human immun
22	2895.5	83.6	861	1 ENV_HV1KB	P31819 human immun
23	2895.5	83.6	868	1 ENV_HV1C4	P05879 human immun
24	2705	78.2	855	1 ENV_HV1Z6	P04580 human immun
25	2695	77.8	863	1 ENV_HV1Z8	P05882 human immun
26	2686	77.6	853	1 ENV_HV1Z2	P12487 human immun
27	2657.5	76.8	859	1 ENV_HV1MA	P04583 human immun
28	2638.5	76.2	856	1 ENV_HV1ZH	P05881 human immun
29	2614.5	75.5	846	1 ENV_HV1ND	P14879 human immun
30	2614	75.5	853	1 ENV_HV1BL	P04581 human immun
31	2244	64.8	854	1 ENV_SIVC2	P17281 chimpanzee
32	1721.5	49.7	421	1 ENV_HV1N5	P12490 human immun
33	1613	46.6	460	1 ENV_HV1Z3	P12491 human immun

ALIGNMENTS

RESULT 1	ID	ENV_HV1JR	STANDARD	PRT	848 AA.	
AC	P20871					Q02837 simian immu
DT	01-FEB-1991 (Rel. 17, Last sequence update)					P20872 human immu
DT	01-FEB-1991 (Rel. 17, Last sequence update)					P18040 human immu
DT	15-SEP-2003 (Rel. 42, Last annotation update)					P27757 simian immu
DE	Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].					P32536 human immu
GN	ENV					P45377 human immu
OS	Human immunodeficiency virus type 1 (JRCSF isolate) (HIV-1).					P27977 simian immu
OC	Viruses; Retroviridae; Retroviridae; Lentiviridae.					P05886 simian immu
OX	NCBI_TaxID=11688;					P24105 human immu
RA	SEQUENCE FROM N.A.					P05883 human immu
RP	Koyanagi S., Chen I.S.Y.;					P17755 human immu
RU	Submitted (DEC-1988) to the HIV data bank.					P12449 human immu
CC	-----					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by, and for commercial entities requires a license agreement (see http://www.ebi.ac.uk/announcements/ or send an email to license@ebi.ac.uk).					
CC	EMBL: M38429; AAB03749.1; ..					
DR	PDB: 1CE4; 1B-MAR-99.					
DR	HIV; M38429; ENV5URCSF.					
DR	InterPro: IPR000328; Env_GP41.					
DR	InterPro: IPR000777; GP120.					
DR	Pfam: PF00516; GP120; 1.					
DR	Pfam: PF00517; GP41; 1.					
KW	AIDS; Coat protein; Polypeptide; Polypeptide; Glycoprotein; Transmembrane; Signal; 3D-structure.					
KW	SIGNAL	1	32			
FT	CHAIN	33	503			EXTERIOR MEMBRANE GLYCOPROTEIN.
FT	CHAIN	504	848			TRANSMEMBRANE GLYCOPROTEIN.
FT	DISULFID	53	73			BY SIMILARITY.
FT	DISULFID	118	203			BY SIMILARITY.
FT	DISULFID	125	194			BY SIMILARITY.
FT	DISULFID	130	154			BY SIMILARITY.
FT	DISULFID	216	245			BY SIMILARITY.
FT	DISULFID	226	237			BY SIMILARITY.
FT	DISULFID	294	328			BY SIMILARITY.
FT	DISULFID	374	437			BY SIMILARITY.
FT	DISULFID	381	410			BY SIMILARITY.
FT	CARBOHYD	87	87			N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	134	134			N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	137	137			N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	153	153			N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	157	157			N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	185	185			N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	195	195			N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	228	228			N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	239	239			N-LINKED (GLCNAC. . .) (POTENTIAL).

```

FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 274 274 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 382 382 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 388 388 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 403 403 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 440 440 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 453 453 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 603 603 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 608 608 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 617 617 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 629 629 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 808 808 N-LINKED (GLCNAC. . .) (POTENTIAL).
SEQUENCE 848 AA; 96475 MM; 2076751227EC3F3 CRC64;

Query Match 90.7%; Score 3139.5; DB 1; Length 848;
Best Local Similarity 91.3%; Pred. No. 1.4e-235;
Matches 589; Conservative 22; Mismatches 31; Indels 3; Gaps 3;

Qy 1 VEKLMVTYYGVPMWKEATTLFCASDAKADTEVHNVMAHACVPTDPNPOEVLENT 60
Db 30 VEKLMVTYYGVPMWKEATTLFCASDAKADTEVHNVMAHACVPTDPNPOEVLENT 89
Qy 61 EHFNMKNKNVNEQOEDIIISLMDQSLKPCVLTPLCVTLNCKDVNATNTTNDSEGTME 120
Db 90 EDFNMKNKNVNEQOEDVINLMDQSLKPCVLTPLCVTLNCKDVNATNTTNDSEGTME 149
Qy 121 EIKKCSFRTTISIDEVQKALFYKLDVVIDN-NTSYLLISCDTSVTYQACPXSFE 179
Db 150 EIKKCSFRTTISIDEVQKALFYKLDVVIDN-NTSYLLISCDTSVTYQACPXSFE 209
Qy 180 PIPHYCAPAGFALILKCNKDTFNGKPCXNVSTQCTGIRPVYSTOLLNGSLAEFEV 239
Db 210 PIPHYCAPAGFALILKCNKDTFNGKPCXNVSTQCTGIRPVYSTOLLNGSLAEFEV 269
Qy 240 IRSDFTNNAKTIIVOLKESVEINCTRENNTRKSHIGPRAFYTTGEIIGDIRAH 299
Db 270 IRSDFTNNAKTIIVOLKESVEINCTRENNTRKSHIGPRAFYTTGEIIGDIRAH 329
Qy 300 ISRAKANDTLKQIYIKAREOFENKTIVNHSFGDPEIWMHSFNCGEFFFCNSTOLF 359
Db 330 ISRAKANDTLKQIYIKAREOFENKTIVNHSFGDPEIWMHSFNCGEFFFCNSTOLF 389
Qy 360 TANNNTGSGNNTEGN-FTTLPCRIKOINMOEYKAMAPPIGQIRCSNITGLL 418
Db 390 TANNNTGSGNNTEGN-FTTLPCRIKOINMOEYKAMAPPIGQIRCSNITGLL 448
Qy 419 DGGINENGTEIFRPGGDMRDNRSEFYKYKVIIEPLGVAPTCKRRVVRKRAVIG 478
Db 449 DGGINENGTEIFRPGGDMRDNRSEFYKYKVIIEPLGVAPTCKRRVVRKRAVIG 508
Qy 479 AVFLGFLGAGSTGASMTLTVQARLLSGIVQOQNLLPAIEAOQMLQITWGIKOL 538
Db 509 AVFLGFLGAGSTGASMTLTVQARLLSGIVQOQNLLPAIEAOQMLQITWGIKOL 568
Qy 539 QARVLAVERYLQDQOLGIMGSGKLTCTTAVPNVNASNSLDRINNMNTMMEBERID 598
Db 569 QARVLAVERYLQDQOLGIMGSGKLTCTTAVPNVNASNSLDRINNMNTMMEBERID 628
Qy 599 NYTSEIYTLIESONOQEKNEOELELDKMASLNMWPDITNMLWY 643
Db 629 NYTSEIYTLIESONOQEKNEOELELDKMASLNMWPDITNMLWY 673

```

```

AC P19550;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein gp160 precursor [contains: Exterior membrane
glycoprotein (gp120); Transmembrane glycoprotein (gp41)].
GN ENV.
OS Human immunodeficiency virus type 1 (SF162 isolate) (HIV-1).
OC Viruses; Retroviridae; Lentivirinae.
OC NCBI_TaxID=11691;
RX [1]
RX MEDLINE=90347835; PubMed=2384920;
RA Cheng-Mayer C., Quirga M., Tung J.W., Levy J.
RT "Viral determinants of human immunodeficiency virus type 1 T-cell or
macrophage tropism, cytopathogenicity, and CD4 antigen modulation.";
RL J. Virol. 64:4390-4398(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M65024; AAA45072.1; -.
CC PDB: 1OBE; 15-MAY-97.
CC HIV: M38428; ENV$SF162.
CC InterPro: IPR000328; Env_Gp41.
CC InterPro: IPR000777; GP120.
CC Pfam: PF00516; GP120; 1.
CC Pfam: PF00517; GP41; 1.
CC AIDS; Coat protein; Glycoprotein; Transmembrane; Signal;
KM 3D-structure.
FT SIGNAL 1 29
FT CHAIN 30 502
FT CHAR 503 847
FT DISULFID 53 73
FT DISULFID 118 203
FT DISULFID 125 194
FT DISULFID 130 155
FT DISULFID 216 245
FT DISULFID 226 237
FT DISULFID 294 328
FT DISULFID 374 435
FT DISULFID 381 408
FT CARBOHYD 87
FT CARBOHYD 135 135
FT CARBOHYD 154 154
FT CARBOHYD 186 186
FT CARBOHYD 195 195
FT CARBOHYD 232 232
FT CARBOHYD 239 239
FT CARBOHYD 260 260
FT CARBOHYD 274 274
FT CARBOHYD 293 293
FT CARBOHYD 299 299
FT CARBOHYD 329 329
FT CARBOHYD 336 336
FT CARBOHYD 352 352
FT CARBOHYD 382 382
FT CARBOHYD 388 388
FT CARBOHYD 392 392
FT CARBOHYD 398 398
FT CARBOHYD 401 401
FT CARBOHYD 438 438
FT CARBOHYD 454 454
FT CARBOHYD 602 602
FT CARBOHYD 607 607
FT CARBOHYD 616 616
FT CARBOHYD 628 628
SEQUENCE 847 AA; 96135 MM; 0A901317DFJF2AB CRC64;

```


Query Match 87.8%; Score 3041; DB 1; Length 847;
 Best Local Similarity 88.7%; Pred. No. 5,9e-228;
 Matches 572; Conservative 29; Mismatches 40; Indels 4; Gaps 4;

QY 1 VEKMTVTVYGVVWKEATTTTFCASDAKAYTEVHNWATHACVPTDPNPOEVLVNT 60
 DB VEKMTVTVYGVVWKEATTTTFCASDAKAYTEVHNWATHACVPTDPNPOEVLVNT 89
 QY 61 EHFNNKNNWVQWOMEDIISLWDSLKPCVKLTPLCVTLNCKDV-NATNTTDSGTER 119
 DB EHFNNKNNWVQWOMEDIISLWDSLKPCVKLTPLCVTLNCKDV-NATNTTDSGTER 149
 QY 90 EHFNNKNNWVQWOMEDIISLWDSLKPCVKLTPLCVTLNCKDV-NATNTTDSGTER 149
 DB EHFNNKNNWVQWOMEDIISLWDSLKPCVKLTPLCVTLNCKDV-NATNTTDSGTER 149
 QY 120 GEIKNSFNITTSIRDEVOKEALFYKLDVPIIDNNNTSYRLISCDTSYITQACPKISFE 179
 DB GEIKNSFNITTSIRDEVOKEALFYKLDVPIIDNNNTSYRLISCDTSYITQACPKISFE 209
 QY 150 GEIKNSFNITTSIRDEVOKEALFYKLDVPIIDNNNTSYRLISCDTSYITQACPKISFE 209
 DB GEIKNSFNITTSIRDEVOKEALFYKLDVPIIDNNNTSYRLISCDTSYITQACPKISFE 209
 QY 180 PIPHYCAPAGFAILKCDKFTNGKGPCKNVSTVOCCTGIRPVSTOLLNGLAEVY 239
 DB PIPHYCAPAGFAILKCDKFTNGKGPCKNVSTVOCCTGIRPVSTOLLNGLAEVY 269
 QY 210 PIPHYCAPAGFAILKCDKFTNGKGPCKNVSTVOCCTGIRPVSTOLLNGLAEVY 269
 DB PIPHYCAPAGFAILKCDKFTNGKGPCKNVSTVOCCTGIRPVSTOLLNGLAEVY 269
 QY 240 IRSDFPTNNAKTIIVQLKESVEINCTRPNNNTKSIHGPGRAPYATGDIIGDIPQACN 299
 DB IRSDFPTNNAKTIIVQLKESVEINCTRPNNNTKSIHGPGRAPYATGDIIGDIPQACN 329
 QY 270 IRSDFPTNNAKTIIVQLKESVEINCTRPNNNTKSIHGPGRAPYATGDIIGDIPQACN 329
 DB IRSDFPTNNAKTIIVQLKESVEINCTRPNNNTKSIHGPGRAPYATGDIIGDIPQACN 329
 QY 300 ISRAKNDTKOIVIKLREOFENKTIIVFNHSSGDPPEIWMHSGEFPYCNSTOLFNS 359
 DB ISRAKNDTKOIVIKLREOFENKTIIVFNHSSGDPPEIWMHSGEFPYCNSTOLFNS 389
 QY 330 ISRAKNDTKOIVIKLREOFENKTIIVFNHSSGDPPEIWMHSGEFPYCNSTOLFNS 389
 DB ISRAKNDTKOIVIKLREOFENKTIIVFNHSSGDPPEIWMHSGEFPYCNSTOLFNS 389
 QY 360 TNNNTTSGSNTTGTITLPCRKOIINMMOEVGKAMAPPIRGOIRCSNITGLLTRD 419
 DB TNNNTTSGSNTTGTITLPCRKOIINMMOEVGKAMAPPIRGOIRCSNITGLLTRD 447
 QY 390 TW-NNTIGPNNNTG-TITLPCRKOIINMMOEVGKAMAPPIRGOIRCSNITGLLTRD 447
 DB TW-NNTIGPNNNTG-TITLPCRKOIINMMOEVGKAMAPPIRGOIRCSNITGLLTRD 447
 QY 420 GGIN-ENGTEIRPBGSDMRDWRSEFYKYKTKIPELVATTKCKRRVVOEKAVGIG 478
 DB GGIN-ENGTEIRPBGSDMRDWRSEFYKYKTKIPELVATTKCKRRVVOEKAVGIG 507
 QY 448 GKEISNTTELFRPGGDMRDWRSEFYKYKTKIPELVATTKCKRRVVOEKAVGIG 507
 DB GKEISNTTELFRPGGDMRDWRSEFYKYKTKIPELVATTKCKRRVVOEKAVGIG 507
 QY 479 AVFLGFLGAGSTMGASMTLTVQARLLSGIVOOONNLRAIEAQORMLQTLWIGIKOL 538
 DB AVFLGFLGAGSTMGASMTLTVQARLLSGIVOOONNLRAIEAQORMLQTLWIGIKOL 567
 QY 508 AMFLGFLGAGSTMGASMTLTVQARLLSGIVOOONNLRAIEAQORMLQTLWIGIKOL 567
 DB AMFLGFLGAGSTMGASMTLTVQARLLSGIVOOONNLRAIEAQORMLQTLWIGIKOL 567
 QY 539 QARVLAVERYLQDQQLGIGWCSGKLICTTAVPNNASWNSKSLDRIWNNMTWEMEREID 598
 DB QARVLAVERYLQDQQLGIGWCSGKLICTTAVPNNASWNSKSLDRIWNNMTWEMEREID 627
 QY 568 QARVLAVERYLQDQQLGIGWCSGKLICTTAVPNNASWNSKSLDRIWNNMTWEMEREID 627
 DB QARVLAVERYLQDQQLGIGWCSGKLICTTAVPNNASWNSKSLDRIWNNMTWEMEREID 627
 QY 599 NTSSEITLIESONQOEKNEBELLEDKMASLWMPDITWMLNY 643
 DB NTSSEITLIESONQOEKNEBELLEDKMASLWMPDITWMLNY 672
 QY 628 NTSSEITLIESONQOEKNEBELLEDKMASLWMPDITWMLNY 672
 DB NTSSEITLIESONQOEKNEBELLEDKMASLWMPDITWMLNY 672

RESULT 3
 ENV_HVISC STANDARD; PRT; 856 AA.
 ID ENV_HVISC STANDARD; PRT; 856 AA.
 AC P05878;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
 GN GN
 OS Human immunodeficiency virus type 1 (SC isolate) (HIV-1).
 OC Viruses; Retrovirus; Retroviridae; Lentivirus.
 NC NCBI_Taxid:11702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88219542; PubMed=3369091;
 RA Guirao C., Guo H.-G., Franchini G., Aldovini A., Collalti E.,
 RA Farrell K., Wong-Staal F., Gallo R.C., Reitz M.S. Jr.;
 RT "Envelope sequences of two new United States HIV-1 isolates."
 RL Virology 154:531-536(1988).
 CC -I- MISCELLANEOUS: THE SC ISOLATE WAS TAKEN FROM AN ARC PATIENT IN
 CC 1984 IN SOUTHERN CALIFORNIA.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL: M17450; NOT_ANNOTATED_CDS.
 DR HIV: M17450; ENVSSC.
 DR InterPro: IPR000328; Env GP41.
 DR InterPro: IPR000777; GP120.
 DR Pfam: PF00516; GP120; 1.
 DR Pfam: PF00517; GP41; 1.
 KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane;
 KW Signal.
 FT SIGNAL 1 29
 FT CHAIN 30 510
 FT CHAIN 511 856
 FT SITE 760 760
 FT DISULFID 53 73
 FT DISULFID 118 206
 FT DISULFID 125 197
 FT DISULFID 130 160
 FT DISULFID 219 247
 FT DISULFID 228 239
 FT DISULFID 296 330
 FT DISULFID 376 439
 FT DISULFID 383 412
 FT CARBOHYD 87 87
 FT CARBOHYD 129 129
 FT CARBOHYD 135 135
 FT CARBOHYD 140 140
 FT CARBOHYD 143 143
 FT CARBOHYD 159 159
 FT CARBOHYD 163 163
 FT CARBOHYD 189 189
 FT CARBOHYD 198 198
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 302 302
 FT CARBOHYD 331 331
 FT CARBOHYD 338 338
 FT CARBOHYD 354 354
 FT CARBOHYD 360 360
 FT CARBOHYD 384 384
 FT CARBOHYD 394 394
 FT CARBOHYD 400 400
 FT CARBOHYD 405 405
 FT CARBOHYD 442 442
 FT CARBOHYD 457 457
 FT CARBOHYD 611 611
 FT CARBOHYD 616 616
 FT CARBOHYD 625 625
 FT CARBOHYD 637 637
 FT CARBOHYD 674 674
 FT CARBOHYD 816 816
 SQ SEQUENCE 856 AA; 97055 MW; DAF4DA600EBA7A08 CMC64;

Query Match 87.4%; Score 3027.5; DB 1; Length 856;
 Best Local Similarity 87.2%; Pred. No. 6.7e-227;
 Matches 571; Conservative 33; Mismatches 34; Indels 17; Gaps 7;

QY 2 EKLMTVTVYGVVWKEATTTTFCASDAKAYTEVHNWATHACVPTDPNPOEVLVNT 61
 DB EKLMTVTVYGVVWKEATTTTFCASDAKAYTEVHNWATHACVPTDPNPOEVLVNT 90
 QY 62 EKLMTVTVYGVVWKEATTTTFCASDAKAYTEVHNWATHACVPTDPNPOEVLVNT 90
 DB EKLMTVTVYGVVWKEATTTTFCASDAKAYTEVHNWATHACVPTDPNPOEVLVNT 90

Db 91 NFNWKNMWEQMHEDIISLMDOSLRKPCVLTPLCVTLNCTNLNDRSTNATNTSSNRG 150
 QY 116 TMRGEIKNGSFNTTTSIRDEVOKEVALFYLDVDPIDNNNTSYRLISCDPSVTOACPK 175
 Db 151 KMEGEMTNSFNITTSIRSKOKEVALFYLDVDPID--NTSTLLNCTNSVITQACPK 208
 QY 176 ISPEPIPIHYCAPAGFAILKCNDRKTFNGKGPCKNVSTVQCTHGIRPVVSTOLLNGSLAE 235
 Db 209 VSEFPIPIHYCA-RMPAILNCKNKKFNGTGTCTVSTVQCTHGIRPVVSTOLLNGSLAE 267
 QY 236 EEVVIRSDNFNNAKTIIVOLKESVEINCTPRNNNTRKSHIGGRAPYTTGEGITGDIRQ 295
 Db 268 EEVVLRSENFNNAKTIIVOLKESVEINCTPRNNNTRKSHIGGRAPYTTGEGITGDIRQ 327
 QY 296 AHGVISSAKMNDTKOYIKYKREOPENKTIVFNHSSGSDPEIWMHSEFCGSEFFYCNSTQ 355
 Db 328 AHGVISSAKMNDTKOYIKYKREOPENKTIVFNHSSGSDPEIWMHSEFCGSEFFYCNSTQ 387
 QY 356 LFNSTNNNTSGSNNTSGN-TITLPCRIKQIINNMOEYKAMVAPPIRGITCSSNTIGL 414
 Db 388 LFSSTW-NGTEGSSNNTGNDTITLPCRIKEIINMMQEVKAMVAPPIRGVCKSSNTIGL 446
 QY 415 LITDGGINENG-----TEIFRPGGDMRDNRSEFFYKTKVKEPLGVAPTKKRRVYQ 469
 Db 447 LITDGGNSKNKSNKNTETIFRPGGDMRDNRSEFFYKTKVKEPLGVAPTKKRRVYQ 506
 QY 470 REKRAVG-IGAVPIGFGAAGSTMGASMTLTVQARLLSGIVQOQNNLLAIEAQOQRL 528
 Db 507 REKRAVGITGMFGLGAGASTMGATSMITLTVQARLLSGIVQOQNNLLAIEAQOQRL 566
 QY 529 QLTWAGIKQIARVLAVERVYLGDOQLGIGWCSGGLICTAVPWNASNKSILDRINNM 588
 Db 567 QLTWAGIKQIARVLAVERVYLRDQQLGIGWCSGGLICTTTPVNTSMNSKSLDKIMNM 626
 QY 589 TMMWERIDNYTSEITVTLIESQNOEKNOELLDKASLMMWPTITWLMY 643
 Db 627 TMMWERIDNYTSLITLIESQNOEKNOELLDKASLMMWPTITWLMY 691
 RESULT 4
 ENV_HV1Y2 STANDARD; PRT; 843 AA.
 ID ENV_HV1Y2 AC P35961;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Envelope polypeptide GP160 precursor (Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)).
 ENV.
 CC Human immunodeficiency virus type 1 (YU-2 isolate) (HIV-1).
 CC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 CC NCBI_TaxID=36377;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93021387; PubMed=1404605;
 RA Li Y., Hui H., Burgess C.J., Price R.W., Sharp P.M., Hahn B.H., Shaw G.M.;
 RA "Complete nucleotide sequence, genome organization, and biological properties of human immunodeficiency virus type 1 in vivo: evidence for limited defectiveness and complementation";
 RT J. Virol. 66:6587-6600(1992).
 RL J. Virol. 66:6587-6600(1992).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL, M93258; -; NOT_ANNOTATED_CDS.
 DR PIR; H44001; H44001.
 DR PDB; 1G9N; 27-DEC-00.

DR InterPro; IPR000328; Env_GP41.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal;
 KW 3D-structure.
 FT SIGNAL 1 29
 FT CHAIN 30 489
 FT CHAIN 490 843
 FT TRANSMEM 738 755
 FT DISULFID 53 73
 FT DISULFID 118 201
 FT DISULFID 125 192
 FT DISULFID 130 155
 FT DISULFID 214 243
 FT DISULFID 224 235
 FT DISULFID 292 326
 FT DISULFID 373 432
 FT DISULFID 380 405
 FT CARBOHYD 87 87
 FT CARBOHYD 129 129
 FT CARBOHYD 135 135
 FT CARBOHYD 138 138
 FT CARBOHYD 154 154
 FT CARBOHYD 158 158
 FT CARBOHYD 184 184
 FT CARBOHYD 193 193
 FT CARBOHYD 230 230
 FT CARBOHYD 237 237
 FT CARBOHYD 258 258
 FT CARBOHYD 272 272
 FT CARBOHYD 285 285
 FT CARBOHYD 291 291
 FT CARBOHYD 297 297
 FT CARBOHYD 327 327
 FT CARBOHYD 351 351
 FT CARBOHYD 381 381
 FT CARBOHYD 389 389
 FT CARBOHYD 395 395
 FT CARBOHYD 400 400
 FT CARBOHYD 435 435
 FT CARBOHYD 450 450
 FT CARBOHYD 598 598
 FT CARBOHYD 603 603
 FT CARBOHYD 612 612
 FT CARBOHYD 624 624
 FT CARBOHYD 803 803
 SQ SEQUENCE 843 AA; 95648 MW; C69D971C918B71 CRC64;
 Query Match 86.7%; Score 3000; DB 1; Length 843;
 Best Local Similarity 87.9%; Pred. No. 8,9e-225;
 Matches 566; Conservative 29; Mismatches 41; Indels 8; Gaps 6;
 QY 2 EKLMTVYVYGVVWKEATTTLLFCASDAKAYDTEVHNVMAHACVPTDPNQEVLNVTE 61
 Db 31 EQLMTVYVYGVVWKEATTTLLFCASDAKAYDTEVHNVMAHACVPTDPNQEVLNVTE 90
 QY 62 HNNWKNMWEQMHEDIISLMDOSLRKPCVLTPLCVTLNCTNLNDRSTNATNTSSNRG 120
 Db 91 NFNWKNMWEQMHEDIISLMDOSLRKPCVLTPLCVTLNCTNLNDRSTNATNTSSNRG 150
 QY 121 EIKKCSFNITTSIRDEVOKEVALFYLDVDPIDNNNTSYRLISCDPSVTOACPKISFEP 180
 Db 151 EIKKCSFNITTSIRDEVOKEVALFYLDVDPID--NASTYLLSNTSVITQACPKVSFEP 208
 QY 181 IPIHYCAPAGFAILKCNDRKTFNGKGPCKNVSTVQCTHGIRPVVSTOLLNGSLAE 240
 Db 209 IPIHYCAPAGFAILKCNDRKTFNGKGPCKNVSTVQCTHGIRPVVSTOLLNGSLAE 268
 QY 241 RSDNFTNNAKTIIVOLKESVEINCTPRNNNTRKSHIGGRAPYTTGEGITGDIRQ 300
 Db 269 RSDNFTNNAKTIIVOLKESVEINCTPRNNNTRKSHIGGRAPYTTGEGITGDIRQ 328

FT	CARBOHYD	239	239	N-LINKED	(GLCNAC . .)	(POTENTIAL) .
FT	CARBOHYD	246	246	N-LINKED	(GLCNAC . .)	(POTENTIAL) .
FT	CARBOHYD	267	267	N-LINKED	(GLCNAC . .)	(POTENTIAL) .
FT	CARBOHYD	281	281	N-LINKED	(GLCNAC . .)	(POTENTIAL) .
FT	CARBOHYD	294	294	N-LINKED	(GLCNAC . .)	(POTENTIAL) .
FT	CARBOHYD	300	300	N-LINKED	(GLCNAC . .)	(POTENTIAL) .
FT	CARBOHYD	306	306	N-LINKED	(GLCNAC . .)	(POTENTIAL) .
FT	CARBOHYD	337	337	N-LINKED	(GLCNAC . .)	(POTENTIAL) .
FT	CARBOHYD	344	344	N-LINKED	(GLCNAC . .)	(POTENTIAL) .
FT	CARBOHYD	361	361	N-LINKED	(GLCNAC . .)	(POTENTIAL) .
FT	CARBOHYD	391	391	N-LINKED	(GLCNAC . .)	(POTENTIAL) .
FT	CARBOHYD	397	397	N-LINKED	(GLCNAC . .)	(POTENTIAL) .
FT	CARBOHYD	402	402	N-LINKED	(GLCNAC . .)	(POTENTIAL) .
FT	CARBOHYD	411	411	N-LINKED	(GLCNAC . .)	(POTENTIAL) .
FT	CARBOHYD	453	453	N-LINKED	(GLCNAC . .)	(POTENTIAL) .
FT	CARBOHYD	468	468	N-LINKED	(GLCNAC . .)	(POTENTIAL) .
FT	CARBOHYD	616	616	N-LINKED	(GLCNAC . .)	(POTENTIAL) .
FT	CARBOHYD	621	621	N-LINKED	(GLCNAC . .)	(POTENTIAL) .
FT	CARBOHYD	630	630	N-LINKED	(GLCNAC . .)	(POTENTIAL) .
FT	CARBOHYD	642	642	N-LINKED	(GLCNAC . .)	(POTENTIAL) .
FT	CARBOHYD	679	679	N-LINKED	(GLCNAC . .)	(POTENTIAL) .
FT	CARBOHYD	755	755	N-LINKED	(GLCNAC . .)	(POTENTIAL) .
FT	CARBOHYD	821	821	N-LINKED	(GLCNAC . .)	(POTENTIAL) .
SQ	SEQUENCE	861 AA;	97487 MM;	04DEB2B4D4EFD63A	CRC64;	
<hr/>						
Query Match		86.4%;	Score 2992.5;	DB 1;	Length 861;	
Beet Local Similarity		86.7%;	Pred. No. 3.5e-224;			
Matches		569;	Conservative	32;	Mismatches	40;
					Indels	15;
					Gaps	8
<hr/>						
OY		2	EKLWVTVYVGVPVWEKATTTLFCASDAKAAYTEVHNVNATHACVPTDENPOEVLENTE	61		
Db		32	EKLWTVYYGVGPVWKERTTLTFCASDAKAYTEVHNWATHACVPTDENPOEVLVNTE	91		
OY		62	HFNWAKNNWVEQOMEDIILSWDOSLKPCVKLTPLCVTLNKDV-NAINT---INDSG-	115		
Db		92	NFNWAKNDWVEQOMEDIILSWDOSLKPCVKLTPLCVSLKCTDLGNAITNSNTSSSGE	151		
OY		116	-TWRGELKNSEKITTSIRDEVOKEVALFKVLDPVPIDNNNTSRLISCSPTSVTAOC	174		
Db		152	MMMKGEIKNCSPNISIRKVKKEVAFFKLDIIPIDNDTSTLTLSCNTSVTAQCP	211		
OY		175	KISEFPIPIHCAAGFAILKCNDKTENGKOPCKNVSVQCETHGRPVASTOLLNGSLA	234		
Db		212	KVSEPEPIPHACAAGFAILLKNNKTNGTGTCPNVSVQCETHGRPVASTOLLNGSLA	271		
OY		235	EEBVIVRSDFNTNAKTIIVOLKESVEINCTRPNNNTRKSJHI--GPGRAFYTGELLGD	292		
Db		272	EEBVIVRSANFTDAKTIIVQLNQSVIEINCTRPNNNTRKSIRIQGPGRAFEVTIGK-IGN	330		
OY		293	IROAHCHISRAXMNDTLKOIYIKLRBP-EKTIIVFNHSSGGDEPIVHNSFNCSEEFYC	351		
Db		331	MROAHCHISRAXMNAATLKQIAASKLRBPFGNNKTIIFKXSSGGDEPIVHSFNCSEEFYC	390		
OY		352	NSTOLFNSTMN--NTEGSNNTEG-NTITLPCRHKOIINMOEGKMVAPPIRGQRIC	407		
Db		391	NSTOLFENSTFNSTWSBTGBSNNTGSDTIIIPCRIKQIINMOEGKMAFPISGQIRC	450		
OY		408	.SSNTTGLLTRDGGINNGTEIFPPGGDMRDNRSEFYKYKIETPLGYAPTCKCARV	467		
Db		451	.SSNTTGLLTRDGGNNNGSGEIFPPGGDMRDNRSELKYKKVYKIETPLGYAPTKARRV	510		
OY		468	VOREKRAVGIGAVTLGFLGAAGTWCQAASMTLTYOARLLLSGIYOQNNLRAIEAOQM	527		
Db		511	VOREKRAVGIGALTLGFLGAAGTWCARSMTLTYOARLLLSGIYOQNNLRAIEAOQHL	570		
OY		528	LQTLVMGIKLOAQLAVERYLGPOOLLGIMGSGSKLICCAVAPNNAWSMSKSIDRIMNN	587		
Db		571	LQTLVMGIKLOAQLAILAVERYLKQOOLLGIMGSGSKICTTAVAPNNAWSMSKSIDRMNN	630		
OY		588	MTMWEMEREIDNYSEIYTTLIEESQNOOEKNEOEHLLEBDKMASLMNWEDINTMLMY	643		
Db		631	MTMWEMERENNYSLSLHSLIEESQNOOEKNEOEHLLEBDKMASLMNWENINTMLMY	686		

[illegible]

RL Nature 313:450-458(1985).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announce/>
 CC or send an email to license@ebi.ac.uk).

DR EMBL, K02083; AAB59873.1; -; ALT_SEQ.
 DR EMBL, X01762; CAA25903.1; ALT_SEQ.
 DR PIR, A03974; VCLJVL.
 DR HIV, K02083; ENVSPV22.
 DR InterPro: IPR000328; Env GP41.
 DR InterPro: IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 DR AIDS; Coat protein; Glycoprotein; Transmembrane;
 DR Signal.

FT CHAIN 1 30 EXTERIOR MEMBRANE GLYCOPROTEIN.
 FT CHAIN 512 856 TRANSMEMBRANE GLYCOPROTEIN.
 FT DISULFID 54 74 BY SIMILARITY.
 FT DISULFID 119 205 BY SIMILARITY.
 FT DISULFID 126 196 BY SIMILARITY.
 FT DISULFID 131 157 BY SIMILARITY.
 FT DISULFID 218 247 BY SIMILARITY.
 FT DISULFID 228 239 BY SIMILARITY.
 FT DISULFID 296 331 BY SIMILARITY.
 FT DISULFID 378 445 BY SIMILARITY.
 FT DISULFID 385 418 BY SIMILARITY.
 FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 156 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 625 625 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 856 AA; 97339 MW; 5FCD81DC3C1209B3 CRC64;

Query Match 85.9%; Score 2974; DB 1; Length 856;
 Best Local Similarity 86.3%; Pred. No. 9.5e-223;
 Matches 562; Conservative 34; Mismatches 45; Indels 10; Gaps 6;

Qy 2 EKWVTVYGVVWVKEATTTTCASDARVYDPEVHNVWATHACVPTDNPQVLENTTE 61
 Db 32 EKWVTVYGVVWVKEATTTTCASDARVYDPEVHNVWATHACVPTDNPQVLENTTE 91

Qy 62 HFNMMKNVNEQOEDIIISMDQSLKPCVLPCLNCKDVANNTNDSG--TMR 119
 Db 92 NFNMMKNVNEQOEDIIISMDQSLKPCVLPCLNCKDVANNTNDSGSGRMIMK 151
 Qy 120 GEIKNGSFNTTTSIRBQVQKVALFYKLDVPEIDNNNTSRLSCDTSVTOACPKISFE 179
 Db 152 GEIKNGSFNTTTSIRBQVQKVALFYKLDVPEIDNNNTSRLSCDTSVTOACPKISFE 211
 Qy 180 PIPHYCAPAFALIKCNDKTFNGKPCCKNVSTVQCCTHGRIPVSTOLLNGSLAESEV 239
 Db 212 PIPHYCAPAFALIKCNDKTFNGKPCCKNVSTVQCCTHGRIPVSTOLLNGSLAESEV 271
 Qy 240 IRSDNFTNNAKTIIVQKESVEINCTRPNNTRKSHI--GGPRAFTTGETIGDIRQAH 297
 Db 272 IRANFTDNAKTIIVQKESVEINCTRPNNTRKSHI--GGPRAFTTGETIGDIRQAH 330
 Qy 298 CNISRAKMNNTLQOYVYKLEQF-EKTYIFNHSSCGDEPVIWHSFNGEPEFYCNSTOL 356
 Db 331 CNISRAKMNNTLQOYVYKLEQF-EKTYIFNHSSCGDEPVIWHSFNGEPEFYCNSTOL 390
 Qy 357 FNTSTNN--NTSGSNNTSG--NTITLPCRKQIINMVEYKAMVAPPIRQIRCSNIT 412
 Db 391 FNTSTNN--NTSGSNNTSG--NTITLPCRKQIINMVEYKAMVAPPIRQIRCSNIT 450
 Qy 413 GLLLTRDGINENGTEIFRPGGDMRDMRSEFYKRVKIEPLGVAIPYKRRVQREK 472
 Db 451 GLLLTRDGINENGTEIFRPGGDMRDMRSEFYKRVKIEPLGVAIPYKRRVQREK 510
 Qy 473 RANGIGAVPFGFGAAGSTMGASMTLVQARLLSGIVQOQNNLRAIPAOQMLQTV 532
 Db 511 RANGIGAVPFGFGAAGSTMGASMTLVQARLLSGIVQOQNNLRAIPAOQMLQTV 570
 Qy 533 WGIKQLQARVAVERYGLDQQLGIVGCSGKLTCTAVPNNASMSNKLDRINNNMTWME 592
 Db 571 WGIKQLQARVAVERYGLDQQLGIVGCSGKLTCTAVPNNASMSNKLDRINNNMTWME 630
 Qy 593 WERIDNTSEIYTLIESQNOQKNEQELLEDKRNASLNNWDTYMLWY 643
 Db 631 WERIDNTSEIYTLIESQNOQKNEQELLEDKRNASLNNWDTYMLWY 681

RESULT 11
 ENV_HVLA2 STANDARD; PRT; 855 AA.
 ID ENV_HVLA2
 AC P03378;
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Envelope glycoprotein GP120 precursor [Contains: Exterior membrane
 glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
 GN ENV.
 OS Human immunodeficiency virus type 1 (ARV2/SF2 isolate) (HIV-1).
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11685;
 RN [1]
 RP MEDLINE=8509453; PubMed=2578227;
 RA Sanchez-Pescador R., Power M.D., Barr P.J., Steimer K.S.,
 RA Stempien M.W., Brown-Shimer S.L., Gee W.W., Renard A., Randolph A.,
 RA Levy J.A., Dina D., Luciw P.A.;
 RT "Nucleotide sequence and expression of an AIDS-associated retrovirus
 (ARV-2)".
 RL Science 227:484-492(1985).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announce/>
 CC or send an email to license@ebi.ac.uk).

DR EMBL, K02007; AAB59882.1; -;

DR PDB; 1K33; 10-OCT-01.
 DR PDB; 1K34; 10-OCT-01.
 DR HIV; K03455; ENVSHX2.
 DR InterPro; IPR000328; Env_GP41.
 DR InterPro; IPR007777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 DR AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal;
 3D-structure.
 KM SIGNAL 1 30
 FT CHAIN 31 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 418
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT CARBOHYD 356 356
 FT CARBOHYD 386 386
 FT CARBOHYD 392 392
 FT CARBOHYD 397 397
 FT CARBOHYD 406 406
 FT CARBOHYD 448 448
 FT CARBOHYD 463 463
 FT CARBOHYD 611 611
 FT CARBOHYD 616 616
 FT CARBOHYD 624 624
 FT CARBOHYD 637 637
 FT CARBOHYD 674 674
 FT CARBOHYD 750 750
 FT CARBOHYD 816 816
 SQ SEQUENCE 856 AA; 97212 MW; 6FAB16AF85107FEO CRC64;
 Query Match 85.6%; Score 2964; DB 1; Length 856;
 Best Local Similarity 86.0%; Pred. No. 5.7e-22;
 Matches 560; Conservative 36; Mismatches 45; Indels 10; Gaps 6;

QY 240 IRSDNFTNNAKTIIVOLKESVEINCRPNNTKRSIHI--GPRAFYTTGEIGDIRQAH 297
 DB 272 IRSVNFTDNAKTIIVOLNTSVEINCRPNNTKRIIRIQGPRAFYTTG-IGNRQAH 330
 QY 298 CNISRAKMNNTLQKQVYIKLEQF-ENKTIIVNHSQGDPEIVNHSFNCGEFFYCNSTOL 356
 DB 331 CNISRAKMNNTLQKQVYIKLEQF-ENKTIIVNHSQGDPEIVNHSFNCGEFFYCNSTOL 390
 QY 357 FNSTMN--NTEGSNTEG-NTTLPKRIKQIINMVEYKAMVAPPIGQIRCSNIT 412
 DB 391 FNSTMN--NTEGSNTEG-NTTLPKRIKQIINMVEYKAMVAPPIGQIRCSNIT 450
 QY 413 GLLTTRDGINNGTEIFRPGGDMNDMSSEFYKYVYKIEPLGVAFTCKRRVQREK 472
 DB 451 GLLTTRDGINNGTEIFRPGGDMNDMSSEFYKYVYKIEPLGVAFTCKRRVQREK 510
 QY 473 RAVGIGAVFLGFGAAGSTWGAASMTLVQARLLSGIVQOONNLLRAIEAQRMQLTV 532
 DB 511 RAVGIGAVFLGFGAAGSTWGAASMTLVQARLLSGIVQOONNLLRAIEAQRMQLTV 570
 QY 533 WGIKQLOARVLAVERVYIGDQLLGIVGCGSKLTCTAVPNAASMSKSLRINNNMTME 592
 DB 571 WGIKQLOARVLAVERVYIGDQLLGIVGCGSKLTCTAVPNAASMSKSLRINNNMTME 630
 QY 593 WEREIDNVTSEIYTLIESONQOEKNEBELLEDKMASLNWPFDTWMLY 643
 DB 631 WEREIDNVTSEIYTLIESONQOEKNEBELLEDKMASLNWPFDTWMLY 681
 RESULT 13
 ENV HV1W1 STANDARD; PRT; 856 AA.
 ID ENV HV1W1
 AC P31872;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
 GN ENV.
 OS Human immunodeficiency virus type 1 (HIV-1).
 OC Viruses; Retrovirus; Retroviridae; Lentivirus.
 OX NCBI_Taxid=31678;
 RN [1]
 RP SEQUENCE FROM N.A.
 EX MEDLINE=66218077; PubMed=2423250.
 RA Steadich B.R., Hahn B.H., Shaw G.M., McNeely P.D., Modrow S., Wolf H., Parks E.S., Parks W.P., Josephs S.F., Gallo R.C., Wong-Staal F.;
 RT "Identification and characterization of conserved and variable regions in the envelope gene of HTLV-III/LAV, the retrovirus of AIDS."
 RT AIDS 45:637-648(1986).
 RL Cell 45:637-648(1986).
 CC -1- MISCELLANEOUS: ISOLATES WMJ1, WMJ2, AND WMJ3 WERE OBTAINED FROM BLOOD SAMPLES SEQUENTIALLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO WAS PERINATALLY INFECTED BY HER MOTHER.
 CC PIR: A24774; VCLJ3W.
 DR PDB; 1LB0; 04-DEC-02.
 DR PDB; 1LCX; 04-DEC-02.
 DR InterPro; IPR000328; Env_GP41.
 DR InterPro; IPR007777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 DR AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal;
 3D-structure.
 KM SIGNAL 1 29
 FT CHAIN 30 510
 FT CHAIN 511 856
 FT DISULFID 53 73
 FT DISULFID 118 205
 FT DISULFID 125 196
 FT DISULFID 130 152
 FT DISULFID 218 247
 FT DISULFID 228 239
 EXTERIOR MEMBRANE GLYCOPROTEIN.
 TRANSMEMBRANE GLYCOPROTEIN.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.

FT	DISULFD	296	440	BY SIMILARITY.
FT	DISULFD	376	444	BY SIMILARITY.
FT	DISULFD	383	417	BY SIMILARITY.
FT	CARBOHYD	87	87	N-LINKED (GLCNAC. .) (POTENTIAL)
FT	CARBOHYD	134	134	N-LINKED (GLCNAC. .) (POTENTIAL)
FT	CARBOHYD	140	140	N-LINKED (GLCNAC. .) (POTENTIAL)
FT	CARBOHYD	151	151	N-LINKED (GLCNAC. .) (POTENTIAL)
FT	CARBOHYD	155	155	N-LINKED (GLCNAC. .) (POTENTIAL)
FT	CARBOHYD	183	183	N-LINKED (GLCNAC. .) (POTENTIAL)
FT	CARBOHYD	197	197	N-LINKED (GLCNAC. .) (POTENTIAL)
FT	CARBOHYD	234	234	N-LINKED (GLCNAC. .) (POTENTIAL)
FT	CARBOHYD	241	241	N-LINKED (GLCNAC. .) (POTENTIAL)
FT	CARBOHYD	262	262	N-LINKED (GLCNAC. .) (POTENTIAL)
FT	CARBOHYD	276	276	N-LINKED (GLCNAC. .) (POTENTIAL)
FT	CARBOHYD	289	289	N-LINKED (GLCNAC. .) (POTENTIAL)
FT	CARBOHYD	295	295	N-LINKED (GLCNAC. .) (POTENTIAL)
FT	CARBOHYD	331	331	N-LINKED (GLCNAC. .) (POTENTIAL)
FT	CARBOHYD	338	338	N-LINKED (GLCNAC. .) (POTENTIAL)
FT	CARBOHYD	354	354	N-LINKED (GLCNAC. .) (POTENTIAL)
FT	CARBOHYD	360	360	N-LINKED (GLCNAC. .) (POTENTIAL)
FT	CARBOHYD	390	390	N-LINKED (GLCNAC. .) (POTENTIAL)
FT	CARBOHYD	394	394	N-LINKED (GLCNAC. .) (POTENTIAL)
FT	CARBOHYD	404	404	N-LINKED (GLCNAC. .) (POTENTIAL)
FT	CARBOHYD	447	447	N-LINKED (GLCNAC. .) (POTENTIAL)
FT	CARBOHYD	459	459	N-LINKED (GLCNAC. .) (POTENTIAL)
FT	CARBOHYD	611	611	N-LINKED (GLCNAC. .) (POTENTIAL)
FT	CARBOHYD	616	616	N-LINKED (GLCNAC. .) (POTENTIAL)
FT	CARBOHYD	625	625	N-LINKED (GLCNAC. .) (POTENTIAL)
FT	CARBOHYD	637	637	N-LINKED (GLCNAC. .) (POTENTIAL)
50	SEQUENCE	856 AA;	97526 MW;	DB6BD1E9C0409B9 CRC64;

Query Match	85.5%;	Score 2958.5;	DB 1;	Length 856;
Best Local Similarity	86.1%;	Pred. No. 1.5e-221;		
Matches 564;	Conservative 30;	Mismatches 46;	Indels 15;	Gaps 7;

```

0Y 1 BEKJWVYVYGPWKEATITLLFCASDARAYPEVHNWMA7HACVPTDPPDCEVLENT 60
Db 30 VEQLWVYVYGPWKEATITLLFCASDARAYPEVHNWMA7HACVPTDPPDCEVLENT 89
0Y 61 EHFNMKNMWEQOMEDIISLMDQSLKPCVKLPLCVTLNCKOVNATNTNDEGTMERG 120
Db 90 ENFNMKNMWEQOMEDIISLMDQSLKPCVKLPLCVTLNCKOVNATNTNDEGTMERG 147
0Y 121 EIKNCSFNITISIRDEVOKEVALFYKLDVVPIDNNN-----TSYRLISCDTSYITQACP 175
Db 148 EYKNCSPNITISIRDEVOKEVALFYKLDVVPIDNNN-----TSYRLISCDTSYITQACP 207
0Y 176 ISFEBPIPHYCAPAGFAIILCKNDKTFNGKGPCKNVSIVOCTHGIRPVSVTOLLNGLAE 235
Db 208 VSFEPIPIHYCAPAGFAIILCKNDKTFNGKGPCKNVSIVOCTHGIRPVSVTOLLNGLAE 267
0Y 236 BEVVRSDNFMTNATTIYQKESVAINCTRPNNNT-RKSIHIGPRAFTTBEIIGDIR 294
Db 268 BEIVRSENFDTNATTIYHLMESVAINCTRPNNNRHRIHIGPRAFTTBEIIGDIR 326
0Y 295 OAHCHISRAKNDTLLKOIYKLRQEPENNTIVNHSGGDPETIVMSFNCESEFPFCNST 354
Db 327 OAHCHISRAKNDTLLKOIYKLRQEPENNTIVNHSGGDPETIVMSFNCESEFPFCNST 386
0Y 355 QLFNSTWN---NTEGSNNTB--GNITLTPCRKQIINNMQEVGKAMYAPRIGQIRCS 409
Db 387 QLFNSTWNVTGISTEGNNTBENGDTITLPCRKQIINNMQEVGKAMYAPRIGQIRCS 446
0Y 410 NITGILLTRDGINENGTETFRPGGDMKMDNRSEFYKYKVYKIEPLGVAFTKCRKRV 469
Db 447 NITGILLTRDGNSSREIETFRPGGDMKMDNRSELYKYKVYKIEPLGVAFTKCRKRV 506
0Y 470 REKRAVG-IGAVFLGFLGAAGSTMGAASTLTYOARLLISGIVQOONNLRAIEAOQRL 528
Db 507 REKRAVGAIIGAMFLGFLGAAGSTMGAASTLTYOARLLISGIVQOONNLRAIEAOQRL 566
0Y 529 QLTWAGIOLARVLAVERYLDQOULLGIWGSCKLICTAVPWNMSXKSIDRIWNNM 588

```

DB	QJTWGICQGLQARVLAVERVLRDQGLDLMGCSGLICTTTPVPMASNSKSMDOJMNMM	626
OY	569 TWMEWEREIDNVTSEIYTYLIEBSONQOEKNEDELLDCKMALSMMWFDITTMWLY	643
DB	627 TWMEWEREIDNVTSLIYNLIEESQOEKNEDELLDCKMALSMMWFSITTMWLY	681
RESULT 14		
ENV_HV1H3	STANDARD;	PRT; 856 AA.
AC	P04624;	
DT	13-AUG-1987 (Rel. 05, Created)	
DT	01-FEB-1996 (Rel. 33, Last sequence update)	
DT	15-SEP-2003 (Rel. 42, Last annotation update)	
DE	Envelope polypeptide Gp160 precursor [Contains: Exterior membrane glycoprotein (Gp120); Transmembrane glycoprotein (Gp41)].	
GN	ENV.	
OS	Human immunodeficiency virus type 1 (HXB3 isolate) (HIV-1).	
OC	Viruses; Retroviridae; Lentivirus.	
OX	NCBI_TaxID=11707;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=85228248; PubMed=2988795;	
RA	Crowl R., Ganguly K., Gordon M., Conroy R., Schaber M., Kramer R., Shaw G.M., Wong-Staal P., Reddy E.P.,	
RT	"HIV-111 env gene products synthesized in E. coli are recognized by antibodies present in the sera of AIDS patients.";	
RL	Cell 41:979-986(1985).	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-slb.ch/announce or send an email to license@isb-slb.ch).	
CC	-----	
DR	EMBL; M14100; AAA4679.1; -.	
DR	PDB; 1JAU; 17-OCT-01.	
DR	PDB; 1JAV; 17-OCT-01.	
DR	HIV; M14100; ENVSHX83.	
DR	InterPro; IPR000328; Env GP41.	
DR	InterPro; IPR007772; GP120.	
DR	Pfam; PF00516; GP120, 1.	
KV	AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal; 3D-structure.	
KW	SIGNAL.	
FT	CHAIN 1 30	
FT	CHAIN 31 511	EXTERIOR MEMBRANE GLYCOPROTEIN.
FT	CHAIN 512 856	TRANSMEMBRANE GLYCOPROTEIN.
FT	DISULFID 119 205	BY SIMILARITY.
FT	DISULFID 126 196	BY SIMILARITY.
FT	DISULFID 131 157	BY SIMILARITY.
FT	DISULFID 218 247	BY SIMILARITY.
FT	DISULFID 228 239	BY SIMILARITY.
FT	DISULFID 296 331	BY SIMILARITY.
FT	DISULFID 378 445	BY SIMILARITY.
FT	DISULFID 385 418	BY SIMILARITY.
FT	CARBOHYD 88	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT	CARBOHYD 136	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT	CARBOHYD 141	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT	CARBOHYD 156 156	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT	CARBOHYD 160 160	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT	CARBOHYD 186 186	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT	CARBOHYD 197 197	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT	CARBOHYD 230 230	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT	CARBOHYD 234 234	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT	CARBOHYD 241 241	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT	CARBOHYD 262 262	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT	CARBOHYD 276 276	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT	CARBOHYD 289 289	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT	CARBOHYD 295 295	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT	CARBOHYD 301 301	N-LINKED (GLCNAC . . .) (POTENTIAL).

Query Match	85.4%	Score 2958	DB 1	Length 856
Local Similarity	86.0%	Pred No. 1.7e-221		
Matches 500	Conservative 33	Mismatches 48	Indels 10	Gaps 6

RESULT 15

DT	15-SEP-2003 (Rel. 42, last annotation update)
DE	Envelope polypeptide GP160 precursor [contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP11)].
GN	ENV.
OS	Human immunodeficiency virus type 1 (SF33 isolate) (HIV-1).
OC	Retroviridae; Retrovirus.
OX	NCBI_TaxID=11690;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=90317906; PubMed=2370688;
RA	York-Higgins D., Cheng-Mayer C., Bauer D., Levy J.A., Dina D.;
RT	"Human immunodeficiency virus type 1 cellular host range,
RT	replication, and cytopathicity are linked to the envelope region of
RL	J. Virol. 64:4016-4020(1990). -----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch). -----
CC	
DR	EMBL; M38427; AAA5067.1; .-
DR	PDB; IMEQ; 1I-AAK-02.
DR	HIV; M38427; ENV\$SFR3.
DR	InterPro; IPR000328; Env_GP41.
DR	InterPro; IPR000777; GP120.
DR	Pfam; PF000516; GP120; 1.
KM	Pfam; PF000517; GP41; 1.
KW	AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal; 3D-structure.
FT	SIGNAL 1 31 BY SIMILARITY.
FT	CHAIN 32 506 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT	DISULFID 507 852 TRANSMEMBRANE GLYCOPROTEIN. BY SIMILARITY.
FT	DISULFID 118 73 BY SIMILARITY.
FT	DISULFID 125 197 BY SIMILARITY.
FT	DISULFID 130 156 BY SIMILARITY.
FT	DISULFID 219 248 BY SIMILARITY.
FT	DISULFID 229 331 BY SIMILARITY.
FT	DISULFID 297 340 BY SIMILARITY.
FT	DISULFID 377 439 BY SIMILARITY.
FT	DISULFID 384 412 BY SIMILARITY.
FT	CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 129 129 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 142 142 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 155 155 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 159 159 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 189 189 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 242 242 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 263 263 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 355 355 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 385 385 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 391 391 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 405 405 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 457 457 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 607 607 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 612 612 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 621 621 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 633 633 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 812 812 N-LINKED (GLCMAC. . .) (POTENTIAL).
 SQ SEQUENCE 852 AA; 96663 MW; EB7BBF8D23C9910D CRC64;

Query Match 85.3%; Score 2954.5; DB 1; Length 852;
 Best Local Similarity 85.6%; Pred. No. 3.1e-221;
 Matches 555; Conservative 33; Mismatches 53; Indels 7; Gaps 5;

```

QY 2 EKLAVTVVYGVPMKEATTTLFCASDPAKAYDTEVHNWMA7HACVPTDPNPOEVLENYTE 61
   |||||
DB 31 EKLAVTVVYGVPMKEATTTLFCASDPAKAYDTEVHNWMA7HACVPTDPNPOEVLENYTE 90
   |||||
QY 62 HENMKNNWVEQWEDIIISLDQSLKPCYKLTPLCVTLNCKDV--NATNTNDSBGTMER 119
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 91 HENMKNNWVEQWEDIIISLDQSLKPCYKLTPLCVTLNCKDV--NATNTNDSBGTMER 150
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 120 GEIKKCSFNTITSIDVQKAYALFYKLDVPI--DNNTSYRLISCDTSVITQACPKIS 177
   |||||
DB 151 GEIKKCSFNTITSIDVQKAYALFYKLDVPI--DNNTSYRLISCDTSVITQACPKIS 210
   |||||
QY 178 FEPIPIHYCAPAGFALIKCNDKTFNGKPCKNVSTVQCTHGIRPVYSTOLLNGSLAEE 237
   |||||
DB 211 FEPIPIHYCAPAGFALIKCNDKTFNGKPCKNVSTVQCTHGIRPVYSTOLLNGSLAEE 270
   |||||
QY 238 VIRSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGPRAFYTTGEIIGDIRAH 297
   |||||
DB 271 VIRSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGPRAFYTTGEIIGDIRAH 330
   |||||
QY 298 CNISRAKNNDTLKOYIKLREOFENKTIYFNHSSGGDPREIVMHSFNCBGFYCNSTOLF 357
   |||||
DB 331 CNISRAKNNDTLKOYIKLREOFENKTIYFNHSSGGDPREIVMHSFNCBGFYCNSTOLF 390
   |||||
QY 358 NSTANNTEGSNNTEGNTITLPCRIKQIINMWOEVKAMYPPIRGQIRCSNITGLLLT 417
   |||||
DB 391 NSTANNTEGSNNTEGNTITLPCRIKQIINMWOEVKAMYPPIRGQIRCSNITGLLLT 449
   |||||
QY 418 RDGGINENG-TEIFRPGGDMRDNRSEFEYKTKVVKIEPLGVAPTKCKRRVVQREKRAVG 476
   |||||
DB 450 RDGGINENG-TEIFRPGGDMRDNRSEFEYKTKVVKIEPLGVAPTKCKRRVVQREKRAVG 509
   |||||
QY 477 -IGAVFLGFLGAAGSTMGAASMTLVQARLLSGIVQOONNLLRAIEAQOBLQLTWGI 535
   |||||
DB 510 -IGAVFLGFLGAAGSTMGAASMTLVQARLLSGIVQOONNLLRAIEAQOBLQLTWGI 569
   |||||
QY 536 KQLQARVLAVERYLDQDLGIMGCSGKLICTTAVPANASMSKSLDRIMNTMTWEMER 595
   |||||
DB 570 KQLQARVLAVERYLDQDLGIMGCSGKLICTTAVPANASMSKSLDRIMNTMTWEMER 629
   |||||
QY 596 EIDNTYSEIYTLIESQNOEKNQOELELDKMASIANNFDITNMLMY 643
   |||||
DB 630 EIDNTYSEIYTLIESQNOEKNQOELELDKMASIANNFDITNMLMY 677
   |||||

```

Search completed: December 12, 2003, 12:30:46
 Job time : 11.7921 secs

THIS PAGE BLANK (US 710)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 12:26:14 ; Search time 33.0726 Seconds
(without alignments)
5017.086 Million cell updates/sec

Title: US-10-032-162-13

Perfect score: 3462

Sequence: 1 VEKLMVTYGVGVPMKEAT.....ELDKWASLWNPEDITWLMY 643

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

1 number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp Vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3424	98.9	847	15	Q75760 human immun
2	3116.5	90.0	848	15	Q74999 human immun
3	3086.5	89.2	852	15	Q92761 human immun
4	3075.5	88.8	854	15	O40222 human immun
5	3073.5	88.8	854	15	O78225 human immun
6	3069.5	88.7	855	15	Q03805 human immun
7	3068	88.6	853	15	Q03811 human immun
8	3063.5	88.5	846	15	Q9PXM7 human immun
9	3059.5	88.4	838	15	Q03806 human immun
10	3048	88.0	863	15	O56110 human immun
11	3045.5	88.0	863	15	Q77989 human immun
12	3044	87.9	854	15	Q9WJW5 human immun
13	3039.5	87.8	860	15	O92762 human immun
14	3039	87.8	861	15	O9YP50 human immun
15	3027	87.4	847	15	Q9WJW5 human immun
16					Q8AR21 human immun

17	3027	87.4	849	15	Q77368 human immun
18	3027	87.4	849	15	Q8Q851 human immun
19	3022.5	87.3	843	15	Q70008 human immun
20	3021	87.3	855	15	O8UL64 human immun
21	3021	87.3	863	15	O42031 human immun
22	3019	87.2	851	15	O56562 human immun
23	3018	87.2	863	15	Q9WJW8 human immun
24	3017.5	87.2	860	15	Q9YP43 human immun
25	3017	87.1	843	15	Q70150 human immun
26	3015	87.1	859	15	Q8Q850 human immun
27	3013	87.0	853	15	O56108 human immun
28	3010	86.9	863	15	Q9WJU4 human immun
29	3009.5	86.9	841	15	Q9QXJ1 human immun
30	3008.5	86.9	843	15	Q9QXK8 human immun
31	3007.5	86.9	852	15	O41883 human immun
32	3007	86.9	853	15	Q9J022 human immun
33	3003.5	86.8	856	15	Q72993 human immun
34	3002.5	86.7	846	15	O8UL63 human immun
35	3002.5	86.7	854	15	O92875 human immun
36	3001	86.7	861	15	O9E527 human immun
37	3000.5	86.7	856	15	O71270 human immun
38	3000	86.7	851	15	O8Q852 human immun
39	3000	86.7	861	15	Q9E523 human immun
40	2999	86.6	854	15	O56566 human immun
41	2999	86.6	855	15	O8UL62 human immun
42	2998	86.6	859	15	O71260 human immun
43	2997.5	86.6	840	15	O8Q2X4 human immun
44	2996	86.5	861	15	O56109 human immun
45	2995.5	86.5	850	15	Q70003 human immun

ALIGNMENTS

RESULT 1

Q75760 PRELIMINARY; PRT; 847 AA.

AC Q75760; 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)

DE Envelope glycoprotein.

GN ENV.

OS Human immunodeficiency virus 1.

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OX NCBI_taxid=11676;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=JRF;

RX MEDLINE=87206194; PubMed=3646751;

RA Koyanagi Y., Miles S., Mitsuhashi R.T., Merrill J.E., Vinters H.V., Chen I.S.;

RT "Dual infection of the central nervous system by AIDS viruses with distinct cellular tropisms.";

RL Science 236:819-822(1987).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=JRF;

RX MEDLINE=91043044; PubMed=2172833;

RA O'Brien W.A., Koyanagi Y., Namazie A., Zhao J.Q., Diagne A., Idler K., Zack J.A., Chen I.S.;

RT "HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120 outside the CD4-binding domain.";

RL Nature 348:69-73(1990).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=JRF;

RX MEDLINE=92092169; PubMed=1684385;

RA Pang S., Vinters H.V., Akashi T., O'Brien W.A., Chen I.S.;

RT "HIV-1 env sequence variation in brain tissue of patients with AIDS-related neurologic disease.";

RL J. Acquir. Immune Defic. Syndr. 4:1082-1092(1991).

RN [4]

RP SEQUENCE FROM N.A.
 RC STRAIN=JREFL;
 RA Pang S., Vinters H.V., Akashi T., O'Brien W.A., Chen I.S.,
 RA Koyanagi Y., Namazie A., Zhao J., Diagne A., Jiler K.;
 RA Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.
 DR EMBL; U63632; AAB05604.1;
 DR InterPro; IPR000328; Env GP41.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00517; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 KM AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
 SQ SEQUENCE 847 AA; 96160 MW; 022D5F24B04FB29F CRC64;

Query Match 98.9%; Score 3424; DB 15; Length 847;
 Best Local Similarity 99.2%; Pred. No. 1,1e-276;
 Matches 638; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 VEKLMVTVYGVVPWKEATTTTLCASDAKAYDEVHNVMAHACVPTDPNQEVLLENVT 60
 |||||
 30 VEKLMVTVYGVVPWKEATTTTLCASDAKAYDEVHNVMAHACVPTDPNQEVLLENVT 89
 |||||

QY 61 EHFMMKNNWVEQOEIIISLMDOSLKPCKVLTPLCTLNCKDVNAATNTTDSGTMERG 120
 |||||
 90 EHFMMKNNWVEQOEIIISLMDOSLKPCKVLTPLCTLNCKDVNAATNTTDSGTMERG 149
 |||||

QY 121 EIKKCSFNITTSIRDEVOKEAYLFYKLDVVPIDNNNTSYRLISCDSVTITQACPKISFE 180
 |||||
 150 EIKKCSFNITTSIRDEVOKEAYLFYKLDVVPIDNNNTSYRLISCDSVTITQACPKISFE 209
 |||||

QY 181 IPIHYCAPAGFALLKCNKDTFNGKPCKNVSTVQCTHGIRPVVSTQLLNGSLAEEVVI 240
 |||||
 210 IPIHYCAPAGFALLKCNKDTFNGKPCKNVSTVQCTHGIRPVVSTQLLNGSLAEEVVI 269
 |||||

QY 241 RSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSHIGRAFYTTGGIIGDIRAHANI 300
 |||||
 270 RSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSHIGRAFYTTGGIIGDIRAHANI 329
 |||||

QY 301 SRAKMDTLKOIVIKLREOFENKTVFNHSSGGDEPEIVNHSFNCGEFFYCNSTOLFNS 360
 |||||
 330 SRAKMDTLKOIVIKLREOFENKTVFNHSSGGDEPEIVNHSFNCGEFFYCNSTOLFNS 389
 |||||

QY 361 WNNTEGSSNTEGNTITLPCRKOIINMVEGKAMVAPPIRGQIRCSNITGLLTRDG 420
 |||||
 390 WNNTEGSSNTEGNTITLPCRKOIINMVEGKAMVAPPIRGQIRCSNITGLLTRDG 449
 |||||

QY 421 GINENGTEIFRRGGDMNDNRSEFYKVKVYKIEPLGVAFTKARRVQREKRAVGIGAV 480
 |||||
 450 GINENGTEIFRRGGDMNDNRSEFYKVKVYKIEPLGVAFTKARRVQREKRAVGIGAV 509
 |||||

QY 481 FLGFLGAGSTMGASMTLTVQARLLLSGIYQOQNLLRAIEAOORMLQITVWGIKOLA 540
 |||||
 510 FLGFLGAGSTMGASMTLTVQARLLLSGIYQOQNLLRAIEAOORMLQITVWGIKOLA 569
 |||||

QY 541 RVLAVERYLDQQLLGIWCSGKLICTTAVPWNASWSNKSIDRIWNNNTWMEWEIDNY 600
 |||||
 570 RVLAVERYLDQQLLGIWCSGKLICTTAVPWNASWSNKSIDRIWNNNTWMEWEIDNY 629
 |||||

QY 601 TSEIYTLIEESQNOEKNEOELELDKNASLNNWDTNNWLMY 643
 |||||
 630 TSEIYTLIEESQNOEKNEOELELDKNASLNNWDTNNWLMY 672
 |||||

RESULT 2
 074999 PRELIMINARY; PRT; 848 AA.
 ID 074999
 AC 074999;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Envelope glycoprotein gp160.
 GN ENV.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OX NCBI_TaxId=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JR-CSF;
 RX MEDLINE=97063128; Pubmed=8906996;
 RA Klasse P.J., Boyd M.T., Weiss R.A., Schulz T.F.;
 RT "Mutations in the vpu, env, and nef genes of a syncytium-inducing
 RT variant of HIV type 1 JR-CSF that infects a range of T cell lines";
 RL AIDS Res. Hum. Retroviruses 12:347-350(1996).
 DR EMBL; U45960; AAB60591.1;
 DR InterPro; IPR000328; Env GP41.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00517; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 KM AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
 SQ SEQUENCE 848 AA; 96433 MW; B240019C8737BC3 CRC64;

Query Match 90.0%; Score 3116.5; DB 15; Length 848;
 Best Local Similarity 90.5%; Pred. No. 1,4e-245;
 Matches 584; Conservative 27; Mismatches 31; Indels 3; Gaps 3;

QY 1 VEKLMVTVYGVVPWKEATTTTLCASDAKAYDEVHNVMAHACVPTDPNQEVLLENVT 60
 |||||
 30 VEKLMVTVYGVVPWKEATTTTLCASDAKAYDEVHNVMAHACVPTDPNQEVLLENVT 89
 |||||

QY 61 EHFMMKNNWVEQOEIIISLMDOSLKPCKVLTPLCTLNCKDVNAATNTTDSGTMERG 120
 |||||
 90 EHFMMKNNWVEQOEIIISLMDOSLKPCKVLTPLCTLNCKDVNAATNTTDSGTMERG 149
 |||||

QY 121 EIKKCSFNITTSIRDEVOKEAYLFYKLDVVPIDN-NTTSYRLISCDSVTITQACPKISFE 179
 |||||
 150 EIKKCSFNITTSIRDEVOKEAYLFYKLDVVPIDN-NTTSYRLISCDSVTITQACPKISFE 209
 |||||

QY 180 PIPHYCAPAGFALLKCNKDTFNGKPCKNVSTVQCTHGIRPVVSTQLLNGSLAEEVVI 239
 |||||
 210 PIPHYCAPAGFALLKCNKDTFNGKPCKNVSTVQCTHGIRPVVSTQLLNGSLAEEVVI 269
 |||||

QY 240 IRSNFTNNAKTIIVOLKESVEINCTRPNNNTRKSHIGRAFYTTGGIIGDIRAHANI 299
 |||||
 270 IRSNFTNNAKTIIVOLKESVEINCTRPNNNTRKSHIGRAFYTTGGIIGDIRAHANI 329
 |||||

QY 300 ISRAKMDTLKOIVIKLREOFENKTVFNHSSGGDEPEIVNHSFNCGEFFYCNSTOLFNS 359
 |||||
 330 ISRAKMDTLKOIVIKLREOFENKTVFNHSSGGDEPEIVNHSFNCGEFFYCNSTOLFNS 389
 |||||

QY 360 TWNNTEGSSNTEG-NTITLPCRKOIINMVEGKAMVAPPIRGQIRCSNITGLLTR 418
 |||||
 390 TW-NDTEKSSGTEBQDTITLPCRKOIINMVEGKAMVAPPIRGQIRCSNITGLLTR 448
 |||||

QY 419 DGINENGTEIFRRGGDMNDNRSEFYKVKVYKIEPLGVAFTKARRVQREKRAVGIG 478
 |||||
 449 DGINENGTEIFRRGGDMNDNRSEFYKVKVYKIEPLGVAFTKARRVQREKRAVGIG 508
 |||||

QY 479 AVFLGFLGAGSTMGASMTLTVQARLLLSGIYQOQNLLRAIEAOORMLQITVWGIKOLA 538
 |||||
 509 AVFLGFLGAGSTMGASMTLTVQARLLLSGIYQOQNLLRAIEAOORMLQITVWGIKOLA 568
 |||||

QY 539 QARVLAVERYLDQQLLGIWCSGKLICTTAVPWNASWSNKSIDRIWNNNTWMEWEID 598
 |||||
 569 QARVLAVERYLDQQLLGIWCSGKLICTTAVPWNASWSNKSIDRIWNNNTWMEWEID 628
 |||||

QY 599 NYTSEIYTLIEESQNOEKNEOELELDKNASLNNWDTNNWLMY 643
 |||||
 629 NYTSEIYTLIEESQNOEKNEOELELDKNASLNNWDTNNWLMY 673
 |||||

RESULT 3

092761 PRELIMINARY; PRT; 852 AA.
 ID 092761, 01-NOV-1998 (TrEMBLrel. 08, Created)
 AC 092761, 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Envelope glycoprotein.
 GN ENV.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OC NCBI_Taxid=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SFMS2.9;
 RX MEDLINE=98178716; PubMed=95319894;
 RA McCutchan F.E., Sanders-Buell E., Salminen M.O., Carr J.K.,
 RA Sheppard W.H.;
 RA "Diversity of the human immunodeficiency virus type 1 envelope
 glycoprotein in San Francisco Men's Health Study participants."
 AIDS Res. Hum. Retroviruses 14:329-337(1998).
 DR EMBL, AF025750; AAC0588.1; -;
 DR InterPro; IPR000328; Env GP41.
 DR InterPro; IPR000173; GAP dhdrogenase.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 DR PROSITE; PS00071; GAPDH; 1.
 DR AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
 SQ SEQUENCE 852 AA; 96545 MW; 3C7780DB0611E617 CRC64;

Query Match 89.2%; Score 3086.5; DB 15; Length 852;
 Best Local Similarity 89.5%; Pred. No. 4,1e-243;

Matches 579; Conservative 25; Mismatches 36; Indels 7; Gaps 5;

QY 3 KLMVTVYGVVPWKEATTTLLFCASDAKAYDTEVHNMATHACVPTDPNPOEVLENT 62
 DB 32 QLMVTVYGVVPWKEATTTLLFCASDAKAYDTEVHNMATHACVPTDPNPOEVLENT 91
 QY 63 FNMKNKNVQEQMEDIIISLMDQSLKPCVLTPLCVTLNCKDV-NATNTTNDSEGMERGE 121
 DB 92 FNMKNKNVQEQMEDIIISLMDQSLKPCVLTPLCVTLNCKDV-NATNTTNDSEGMERGE 151
 QY 122 IKNGSFNTTISRDEYQKAYALFYKLDVVPID--NNNTSYRLISCDTSYITQACKRISFE 179
 DB 152 IKNGSFNTTISRDEYQKAYALFYKLDVVPIDNTNTSYRLVSCNTSYITQACKRISFE 211
 QY 180 PIPHYCAPAGFAILKCKNDKTFNGKPCKNVSTVCTGIRPVVSTQULLNGSLAEEVY 239
 DB 212 PIPHYCAPAGFAILKCKNDKTFNGKPCKNVSTVCTGIRPVVSTQULLNGSLAEEVY 271
 QY 240 IRSDFNTNAKTIIVQLKESVEINCTRPNNNTKRSIHIGPRAFYTTGIIIGDIRQACHN 299
 DB 272 IRSENIITNAKTIIVQLKESVEINCTRPNNNTKRSIHIGPRAFYTTGIIIGDIRQACHN 331
 QY 300 ISRAKMDTLKQIVKLEPQENKTVFHHSSGDBEIVMHSFNGEGEFYCNSTOLPNS 359
 DB 332 ISRAKMDTLKQIVKLEPQENKTVFHHSSGDBEIVMHSFNGEGEFYCNSTOLPNS 391
 QY 360 TWMNNTGSGNTEGN--TITLPCRIKQIIMNMQEVKAMYPPIRGQIRCSNITGLLT 417
 DB 392 TW-NTEGSGNTEGN--TITLPCRIKQIIMNMQEVKAMYPPIRGQIRCSNITGLLT 450
 QY 418 RDGGINENG-TETFPFGGDMKDNRRSEFYKXKVVYKIEPLGVAFTCKKRRVYQREKRAVG 476
 DB 451 RDGGINENG-TETFPFGGDMKDNRRSEFYKXKVVYKIEPLGVAFTCKKRRVYQREKRAVG 510
 QY 477 IGAVLGLFGAAGSMGASMTLTVOARLLSGIYQOQNNILRAIEAOQRMQLTVWGK 536
 DB 511 IGAVLGLFGAAGSMGASMTLTVOARLLSGIYQOQNNILRAIEAOQRMQLTVWGK 570
 QY 537 QLOARVLAVERYLQDQLLIGWCSGKLICTAVPMNASWNSKSLDRIVNNNTMMEWERE 596

DB 571 QLOARVLAVERYLQDQLLIGWCSGKLICTTVPNTNSKSKSLDTIINNNTMMEWERE 630

QY 597 IDNTYSEIYTLIEESQNOEKNEQELBLDKKASIMNPDITNMLY 643
 DB 631 IDNTYSEIYTLIEESQNOEKNEQELBLDKKASIMNPDITNMLY 677

RESULT 4

040222 PRELIMINARY; PRT; 854 AA.
 ID 040222, 01-JAN-1998 (TrEMBLrel. 05, Created)
 AC 040222, 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Env polyprotein.
 GN ENV.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OC NCBI_Taxid=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AD8;
 RX MEDLINE=96432129; PubMed=8835195;
 RA Theodore T.S., Englund G., Buckler-White A., Buckler C.E.,
 RA Martin M.A., Peden K.W.;
 RA "Construction and characterization of a stable full-length macrophage-
 RT tropic HIV type 1 molecular clone that directs the production of high
 RT titers of progeny virions."
 RT AIDS Res. Hum. Retroviruses 12:191-194(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AD8;
 RA Theodore T.S., Englund G., Buckler-White A., Buckler C.E.,
 RA Martin M.A., Peden K.W.C.;
 RA Submitted (May-1997) to the EMBL/Genbank/DBJ databases.
 DR EMBL, AF004394; AAB64170.1; -;
 DR InterPro; IPR000328; Env GP41.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 DR AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
 SQ SEQUENCE 854 AA; 97291 MW; 06C45B691036C12 CRC64;

Query Match 88.8%; Score 3075.5; DB 15; Length 854;
 Best Local Similarity 89.4%; Pred. No. 3.3e-242;

Matches 583; Conservative 19; Mismatches 39; Indels 11; Gaps 7;

QY 1 VEKLVTVYGVVPWKEATTTLLFCASDAKAYDTEVHNMATHACVPTDPNPOEVLENT 60
 DB 30 VENLVTVYGVVPWKEATTTLLFCASDAKAYDTEVHNMATHACVPTDPNPOEVLENT 89
 QY 61 EHFNMKNKNVQEQMEDIIISLMDQSLKPCVLTPLCVTLNCKDV-NATNTTNDSEGMERGE 119
 DB 90 EHFNMKNKNVQEQMEDIIISLMDQSLKPCVLTPLCVTLNCKDV-NATNTTNDSEGMERGE 147
 QY 120 GEIKNSFNITTSIRDEYQKAYALFYKLDVVPIDNNNTSYRLISCDTSYITQACKRISFE 179
 DB 148 GEIKNSFNITTSIRDEYQKAYALFYKLDVVPIDNNNTSYRLISCDTSYITQACKRISFE 207
 QY 180 PIPHYCAPAGFAILKCKNDKTFNGKPCKNVSTVCTGIRPVVSTQULLNGSLAEEVY 239
 DB 208 PIPHYCAPAGFAILKCKNDKTFNGKPCKNVSTVCTGIRPVVSTQULLNGSLAEEVY 267
 QY 240 IRSDFNTNAKTIIVQLKESVEINCTRPNNNTKRSIHIGPRAFYTTGIIIGDIRQACHN 299
 DB 268 IRSDFNTNAKTIIVQLKESVEINCTRPNNNTKRSIHIGPRAFYTTGIIIGDIRQACHN 327
 QY 300 ISRAKMDTLKQIVKLEPQENKTVFHHSSGDBEIVMHSFNGEGEFYCNSTOLPNS 358
 DB 338 ISRTKMNNTLQIARKLEQEGNNKTIYFNOSGGDBEIVMHSFNGEGEFYCNSTOLPNS 387
 QY 359 STMNNTGSGNTEGN--TITLPCRIKQIIMNMQEVKAMYPPIRGQIRCSNITGLT 413

```

Db 388 STANFNGTANLTQSGTNGTNDITLPCRIKOIIMMVGKAMAPPIRGIRCSNITG 447
Qy 414 LLLRDPGGINENG-TEIRPRGGGMDRMRSEFYKVKVKEIEPLGVAFTKCRVVOREK 472
Db 448 LILTRDGGNNNNNDTEFRPGGDMRDNRSELYKVKVKEIEPLGVAFTKCRVVOREK 507
Qy 473 RAVG-IGAVFLGFLGAAGSTMGAAASMTLTVQARLLLSGIYQOONNLRAIEAOOQRLQTL 531
Db 508 RAVGTIGMFLGFLGAAGSTMGAAASMTLTVQARLLLSGIYQOONNLRAIEAOOQRLQTL 567
Qy 532 VVGIGLQARVLAVERLYGDOOLIGMGCSGKLICTTAVPWNASWSKSLDRIMNNMTWM 591
Db 568 VVGIGLQARVLAVERLYRDOOLIGMGCSGKLICTTAVPWNASWSKSLDRIMNNMTWM 627
Qy 592 EMEEINNTSEIYTLIEESONOEKNEOELELDKASIMNPDITNMLMY 643
Db 628 EMEEIDNYTGLIYTLIEESONOEKNEOELELDKASIMNPDITNMLMY 679

RESULT 5
Qy 078225 PRELIMINARY; PRT; 854 AA.
AC 078225;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Env protein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ba-1;
RA MEDLINE=91289160; PubMed=1905842;
RA Hwang S.S., Boyle T.J., Lyerly H.K., Cullen B.R.;
RT "Identification of the envelope V3 loop as the primary determinant of
RL cell tropism in HIV-1."
RL Science 253:71-74(1992);
DR EMBL; M63929; AAA75116.1; -
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KM AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 854 AA; 97186 MW; 005FC81B5CBDF8 CRC64;

Query Match 88.8%; Score 3073.5; DB 15; Length 854;
Best Local Similarity 88.6%; Pred. No. 4.7e-242;
Matches 575; Conservative 24; Mismatches 43; Indels 7; Gaps 2;

Qy 2 EKLWVTVYGVVWKEATTTTLCASDAKAYDTEVHNWMAHACVPTDNPQEVLENTTE 61
Db 31 EKLWVTVYGVVWKEATTTTLCASDAKAYDTEVHNWMAHACVPTDNPQEVLENTTE 90
Qy 62 HFNMMKNMVEOMEDIIISLDOSLKPCVKLTPLCVTLNCKDV-----NATNTNDSG 115
Db 91 NFNMMKNMVEOMEDIIISLDOSLKPCVKLTPLCVTLNCKDV-----NATNTNDSG 150
Qy 116 TMRGEIKNGSFNTTTSIRDEVOKEYALFYKLDVVPIDNN-NTSYRLISCDTSVITQACP 174
Db 151 MVGGEKMGKSPKLTITNRGVQKEYALFYELDIPIIDNNSNNRYRLISCDTSVITQACP 210
Qy 175 KISFEPIPIHICAPAGFALLKCKDKKFKNGKPCSNVSTVQCTHGIRPVVSTOLLNLSLA 234
Db 211 KISFEPIPIHICAPAGFALLKCKDKKFKNGKPCSNVSTVQCTHGIRPVVSTOLLNLSLA 270
Qy 235 EEEVVISDNFTNNAKTIIVQLKESVEINCRPNNTKSHIHGGRAFYTTGELIGDIR 294
Db 271 EEEVVISDNFTNNAKTIIVQLKESVEINCRPNNTKSHIHGGRAFYTTGELIGDIR 330
Qy 295 QAHGNIISRAKWNIDLKQIVIKLRQENKTIIVFNHSSGSDPEIVMHSFNGEGEPFYCNST 354

```

```

Db 331 QAHGNIISRAKWNIDLKQIVIKLRQENKTIIVFNHSSGSDPEIVMHSFNGEGEPFYCNST 390
Qy 355 QAHGNIISRAKWNIDLKQIVIKLRQENKTIIVFNHSSGSDPEIVMHSFNGEGEPFYCNST 414
Db 391 QAHGNIISRAKWNIDLKQIVIKLRQENKTIIVFNHSSGSDPEIVMHSFNGEGEPFYCNST 450
Qy 415 LILTRDGGNNNNNDTEFRPGGDMRDNRSELYKVKVKEIEPLGVAFTKCRVVOREK 474
Db 451 LILTRDGGNNNNNDTEFRPGGDMRDNRSELYKVKVKEIEPLGVAFTKCRVVOREK 510
Qy 475 VVGIGLQARVLAVERLYGDOOLIGMGCSGKLICTTAVPWNASWSKSLDRIMNNMTWM 534
Db 511 VVGIGLQARVLAVERLYRDOOLIGMGCSGKLICTTAVPWNASWSKSLDRIMNNMTWM 570
Qy 535 IKQIQAARVLAVERLYGDOOLIGMGCSGKLICTTAVPWNASWSKSLDRIMNNMTWM 594
Db 571 IKQIQAARVLAVERLYRDOOLIGMGCSGKLICTTAVPWNASWSKSLDRIMNNMTWM 630

RESULT 6
Qy 003805 PRELIMINARY; PRT; 855 AA.
AC 003805;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Env protein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC Beaver B.; Popovic M., Gartner S., Gallo R.C., Reed-Connole E.,
RA Submitted (JUN-1991) to the EMBL/GenBank/DBJ databases.
RL EMBL; M68893; AAA44191.1; -
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KM AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 855 AA; 96988 MW; BC631A3B84180CD8 CRC64;

Query Match 88.7%; Score 3069.5; DB 15; Length 855;
Best Local Similarity 88.1%; Pred. No. 1e-241;
Matches 572; Conservative 27; Mismatches 43; Indels 7; Gaps 2;

Qy 2 EKLWVTVYGVVWKEATTTTLCASDAKAYDTEVHNWMAHACVPTDNPQEVLENTTE 61
Db 31 EKLWVTVYGVVWKEATTTTLCASDAKAYDTEVHNWMAHACVPTDNPQEVLENTTE 90
Qy 62 HFNMMKNMVEOMEDIIISLDOSLKPCVKLTPLCVTLNCKDV-----NATNTNDSG 115
Db 91 NFNMMKNMVEOMEDIIISLDOSLKPCVKLTPLCVTLNCKDV-----NATNTNDSG 150
Qy 116 TMRGEIKNGSFNTTTSIRDEVOKEYALFYKLDVVPIDNN-NTSYRLISCDTSVITQACP 174
Db 151 MVGGEKMGKSPKLTITNRGVQKEYALFYELDIPIIDNNSNNRYRLISCDTSVITQACP 210
Qy 175 KISFEPIPIHICAPAGFALLKCKDKKFKNGKPCSNVSTVQCTHGIRPVVSTOLLNLSLA 234
Db 211 KISFEPIPIHICAPAGFALLKCKDKKFKNGKPCSNVSTVQCTHGIRPVVSTOLLNLSLA 270
Qy 235 EEEVVISDNFTNNAKTIIVQLKESVEINCRPNNTKSHIHGGRAFYTTGELIGDIR 294
Db 271 EEEVVISDNFTNNAKTIIVQLKESVEINCRPNNTKSHIHGGRAFYTTGELIGDIR 330
Qy 295 QAHGNIISRAKWNIDLKQIVIKLRQENKTIIVFNHSSGSDPEIVMHSFNGEGEPFYCNST 354

```

```

Db      331 QAHCHLSRAKNDTINKVITKIREQFGNKTIVFKHSSGGDPRIYVHSFPGCEPFYCNSIT 390
Qy      355 QLFNSTNNNTGSGNNTEGNTITLPCRIRKQIINMGOEVKAMYPPIRGQIRCSSNTIGL 414
Db      391 QLFNSTNNNTGSGNNTEGNTITLPCRIRKQIINMGOEVKAMYPPIRGQIRCSSNTIGL 450
Qy      415 LITRDGGINENGTEIFRPGGDMRDNRSEFYKRVKIEPLGVAPTKCRVRVOREKRA 474
Db      451 LITRDGCEPDNKTVEFRPGGDMRDNRSELYKRVKIEPLGVAPTKCRVRVOREKRA 510
Qy      475 VGIGAVFLGFLGAAGSTWGAASNTLTVQARLLLSGIVQOONNLRAIEAOQMLQTLTWG 534
Db      511 VGIGAVFLGFLGAAGSTWGAASNTLTVQARLLLSGIVQOONNLRAIEAOQMLQTLTWG 570
Qy      535 IKOLQARVLAVERYLADQDLGIWCGSGKLICTTAVPMAASMSKSLDRINNNMTWME 594
Db      571 IKOLQARVLAVERYLADQDLGIWCGSGKLICTTAVPMAASMSKSLDRINNNMTWME 630
Qy      595 REIDNYTSEIYTLIEESONQOEKNEQELLELDKMASLNMWFDITNMLWY 643
Db      631 REINNYTSIYSLIESONQOEKNEQELLELDKMASLNMWFDITNMLWY 679

```

RESULT 7

```

ID      003811 PRELIMINARY; PRT; 853 AA.
AC      003811;
DT      01-NOV-1996 (TREMBLrel. 01, Created)
DT      01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT      01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE      Surface envelope glycoprotein.
OS      Human immunodeficiency virus 1.
OC      Viruses; Retroid viruses; Retroviridae; Lentivirus.
RX      NCBI_TaxID=11676;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=ADA;
RX      MEDLINE=91195299; PubMed=2014229;
RA      Westervelt P.; Gendelman H.E.; Ratner L.;
RT      "Identification of a determinant within the human immunodeficiency
RT      virus 1 surface envelope glycoprotein critical for productive
RT      infection of primary monocytes."
RL      Proc. Natl. Acad. Sci. U.S.A. 88:3097-3101(1991).
DR      EMBL; M60472; AAA5065.1; -.
DR      InterPro; IPR000328; Env GP41.
DR      InterPro; IPR000777; GP120.
DR      Pfam; PF00516; GP120; 1.
DR      Pfam; PF00517; GP41; 1.
SQ      Envelope protein.
SEQUENCE 853 AA; 96998 MW; AD2AF21E2B06AD78 CRC64;

```

```

Query Match      88.6%; Score 3068; DB 15; Length 853;
Best Local Similarity 88.8%; Pred. No. 1.3e-241;
Matches 578; Conservative 25; Mismatches 38; Indels 10; Gaps 6;

Qy      1 VEKLVNTYVYGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENT 60
Db      30 VENTLVNTYVYGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENT 89
Qy      61 EHFNMKNMNVQOMQEDIIISLWDQSLKPCVKLTPLCVTLNCKDV-NATNTTDSGTEWER 119
Db      90 ENFMNKNMNVQOMQEDIIISLWDQSLKPCVKLTPLCVTLNCKDLNATNINSSSGM--R 147
Qy      120 GEIKKCSNITTSIRDEVOKEAYALFYKLDVVPIDNNNTSYRLISCDTSVITQACPKISFE 179
Db      148 GEIKKCSNITTSIRDKVKQYALFYRLDVPIDNDNTSYRLINCNTSTITQACPKVSEFE 207
Qy      180 PIPHYCAPAGFAILKCKNDKTFNGKGPCKNVSTVQCTGIRPVVSTOLLNGLSLAESEVY 239
Db      208 PIPHYCTPAGFAILKCKDKKFKNGTGPKCKNVSTVQCTGIRPVVSTOLLNGLSLAESEVY 267
Qy      240 IRSDFNTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPRAFYTTGEIIGDIIQAHCN 299

```

```

Db      268 IRSSNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPRAFYTTGEIIGDIIQAHCN 327
Qy      300 ISRAKMNLTQKQIVTKLEOF-ENKTIYFNHSSGGDPRIVHHSFPGCEPFYCNSITQFN 358
Db      328 ISRTKMNLTQKQIVTKLEOFGNNTIYFNHSSGGDPRIVHHSFPGCEPFYCNSITQFN 387
Qy      359 STWNN---NTEGNNTEGN-TITLPCRIRKQIINMGOEVKAMYPPIRGQIRCSSNTIG 413
Db      388 STWNNNGTNWNLQNSGTGNTITLPCRIRKQIINMGOEVKAMYPPIRGQIRCSSNTIG 447
Qy      414 LITRDGGINENGTEIFRPGGDMRDNRSEFYKRVKIEPLGVAPTKCRVRVOREKRA 473
Db      448 LITRDGCTNNGSGSEIFRPGGDMRDNRSELYKRVKIEPLGVAPTKCRVRVOREKRA 507
Qy      474 AVG-IGAVFLGFLGAAGSTWGAASNTLTVQARLLLSGIVQOONNLRAIEAOQMLQTLTW 532
Db      508 AVGTIGAVFLGFLGAAGSTWGAASNTLTVQARLLLSGIVQOONNLRAIEAOQMLQTLTW 567
Qy      533 WGIKOLQARVLAVERYLADQDLGIWCGSGKLICTTAVPMAASMSKSLDRINNNMTWME 592
Db      568 WGIKOLQARVLAVERYLADQDLGIWCGSGKLICTTAVPMAASMSKSLDRINNNMTWME 627
Qy      593 WEREIDNYTSEIYTLIEESONQOEKNEQELLELDKMASLNMWFDITNMLWY 643
Db      628 WEREIDNYTGLIYTLIEESONQOEKNEQELLELDKMASLNMWFDISNMLWY 678

```

RESULT 8

```

ID      09PXW7 PRELIMINARY; PRT; 846 AA.
AC      09PXW7;
DT      01-MAY-2000 (TREMBLrel. 13, Created)
DT      01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT      01-OCT-2002 (TREMBLrel. 22; Last annotation update)
DE      GP160, envelope glycoprotein.
GN      ENV.
OS      Human immunodeficiency virus 1.
OC      Viruses; Retroid viruses; Retroviridae; Lentivirus.
RX      NCBI_TaxID=11676;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      MEDLINE=93189881; PubMed=8446773;
RA      Schmidtmayerova H.; Gayet O.; Guettari N.; Bolmont C.; Hirsch I.;
RT      "Characterization of HIV1-PAR, a macrophage-tropic strain: cell
RT      tropism, virus/cell entry and nucleotide sequence of the envelope
RT      glycoprotein."
RL      Res. Virol. 144:21-26(1993).
DR      InterPro; IPR000328; Env GP41.
DR      InterPro; IPR000777; GP120.
DR      Pfam; PF00516; GP120; 1.
DR      Pfam; PF00517; GP41; 1.
DR      Pfam; PF00517; GP41; 1.
SQ      AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SEQUENCE 846 AA; 96490 MW; 3D561D5735F90C84 CRC64;

```

```

Query Match      88.5%; Score 3063.5; DB 15; Length 846;
Best Local Similarity 89.3%; Pred. No. 3.1e-241;
Matches 575; Conservative 29; Mismatches 35; Indels 5; Gaps 5;

Qy      2 EKLWNTYVYGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENT 61.
Db      31 DKLWNTYVYGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENT 90
Qy      62 EHFNMKNMNVQOMQEDIIISLWDQSLKPCVKLTPLCVTLNCKDV-NATNTTDSGTEWERG 120
Db      91 YFNFMKNMNVQOMQEDIIISLWDQSLKPCVKLTPLCVTLNCKDLNATNINSSSGM--R 150
Qy      121 EIKKCSNITTSIRDEVOKEAYALFYKLDVVPIDNNNTSYRLISCDTSVITQACPKISFE 180
Db      151 EIKKCSNITTSIRDKVKQYALFYRLDVPIDNNNTSYRLISCDTSVITQACPKVSEFE 210
Qy      181 IPIHYCAPAGFAILKCKNDKTFNGKGPCKNVSTVQCTGIRPVVSTOLLNGLSLAESEVY 240

```

```

Db 211 IPIHYPAGAFALIKCKDKKFDGKPCPNVSTVQCTHGIRPVVSTQLLNGSLAEBEVIR 270
Qy 241 RSDNFTNNAKTIIVQLKESVINCTRPNNTRKSIHIGPGAFTTGGIIGDIOAHONI 300
Db 271 RSDNFTNNAKTIIVQLNESVEINCTRPNNTRKSIHIGPGAFTTGGIIGDIOAHONI 330
Qy 301 SRKANDTLKQIVIKLREOFENKTIIVFNHSGGPEIYVHSGEGEFYCNSTOLFNST 360
Db 331 SKTMENTFOIYKLRQYKNTIIVFNHSGGPEIYVHSGEGEFYCNSTOLFNST 390
Qy 361 MNNTGSSNNTGNTITLPCRICKOIIINMOEYKAMVAPPIRGQIRCSNITGLLTRDG 420
Db 391 W-NTGESSNTE-RTITLPCRICKOIIINMOEYKAMVAPPIRGQIRCSNITGLLTRDG 448
Qy 421 G-INENGTIFRPGGDMRDNRSSEFYKVKYKIEPLGVAFTKCRKRVVQREKAVGIGA 479
Db 449 GNNNNNGEILFRPVGDMRDNRSSEFYKVKYKIEPLGVAFTKCRKRVVQREKAVGIGA 508
Qy 480 VFLGFLGAAGSTMGASMTLTVQARLLLSGIYQOQNNLRAIEAOQRMQLTVMGIKOLQ 539
Db 509 MFLGFLGAAGSTMGARRTITLVQARLLLSGIYQOQNNLRAIEAOQHLQLTVMGIKOLQ 568
Qy 540 ARVLAVERYLDQQLIGWCSGKLICTTAVPNNASWSNKSIDRIINNNTWMEWEEREIDN 599
Db 569 ARVLAVERYLDQQLIGWCSGKLICTTAVPNNASWSNKSIDRIINNNTWMEWEEREIDN 627
Qy 600 YTSIYTLIEESONQOEKNEOELBELDKWASLMMFDTITNMLWY 643
Db 628 YTNLIYTLIEESONQOEKNEOELBELDKWASLMMFDTITNMLWY 671

```

RESULT 9

```

ID 003806 PRELIMINARY; PRT; 838 AA.
AC 003806;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Env protein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_Taxid=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Reitz M., Popovic M., Gartner S., Gallo R.C., Reed-Connole E.,
   Beaver B.;
   Submitted (JUN-1991) to the EMBL/Genbank/DBJ databases.
DR EMBL: M68894; AAA44196.1; -.
DR InterPro: IPR000328; Env GP41.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00517; GP120; 1.
DR Pfam: PF00517; GP41; 1.
SQ SEQUENCE 838 AA; 95329 MW; 11044B1CA10CFB4D CRC64;

```

Query Match 88.4%; Score 3059.5; DB 15; Length 838;

Best Local Similarity 88.8%; Pred. No. 6,4e-241; Indels 9; Gaps 1;

```

Matches 570; Conservative 25; Mismatches 38; Indels 9; Gaps 1;
Qy 2 EKLWTVYVYGVVWKEATTTLFCASDAKAYDEVHNVATHACVPTDPNPOEVLLENVTE 61
Db 31 EKLWTVYVYGVVWKEATTTLFCASDAKAYDEVHNVATHACVPTDPNPOEVLLENVTE 90
Qy 62 HFNMMKNMVEOMQEDIIISLDOSLKPCKVLTPLCVTLNCKDVNATNTTNDSEGMERGE 121
Db 91 NFNMMKNMVEOMQEDIIISLDOSLKPCKVLTPLCVTLNCKDVNATNTTNDSEGMERGE 141
Qy 122 IKNSFNTTISRUBVOKEVALFYKLDVVPIDNNNTSVRLISCDTSVITQACPKISFEPI 181
Db 142 MKNSFNTTISRUBVOKEVALFYKLDVVPIDNNNTSVRLISCDTSVITQACPKISFEPI 201
Qy 182 PIHYPAGAFALIKCKDKKFDGKPCPNVSTVQCTHGIRPVVSTQLLNGSLAEBEVIR 241

```

```

Db 202 PIHYPAGAFALIKCKDKKFDGKPCPNVSTVQCTHGIRPVVSTQLLNGSLAEBEVIR 261
Qy 242 SDNFTNNAKTIIVQLKESVINCTRPNNTRKSIHIGPGAFTTGGIIGDIOAHONI 301
Db 262 SENFTNNAKTIIVQLNESVEINCTRPNNTRKSIHIGPGAFTTGGIIGDIOAHONI 321
Qy 302 RAKANDTLKQIVIKLREOFENKTIIVFNHSGGPEIYVHSGEGEFYCNSTOLFNST 361
Db 332 RAKANDTLKQIVIKLRQYKNTIIVFNHSGGPEIYVHSGEGEFYCNSTOLFNST 391
Qy 362 MNNTGSSNNTGNTITLPCRICKOIIINMOEYKAMVAPPIRGQIRCSNITGLLTRDG 421
Db 382 NTEESNNTENNTITLPCRICKOIIINMOEYKAMVAPPIRGQIRCSNITGLLTRDG 441
Qy 422 INENGTIFRPGGDMRDNRSSEFYKVKYKIEPLGVAFTKCRKRVVQREKAVGIGA 481
Db 442 PEDKTEVFPPGGDMRDNRSSEFYKVKYKIEPLGVAFTKCRKRVVQREKAVGIGA 501
Qy 482 LGFLGAAGSTMGASMTLTVQARLLLSGIYQOQNNLRAIEAOQRMQLTVMGIKOLQ 541
Db 502 LGFLGAAGSTMGASMTLTVQARLLLSGIYQOQNNLRAIEAOQHLQLTVMGIKOLQ 561
Qy 542 VLAVERYLDQQLIGWCSGKLICTTAVPNNASWSNKSIDRIINNNTWMEWEEREIDN 601
Db 562 VLAVERYLDQQLIGWCSGKLICTTAVPNNASWSNKSIDRIINNNTWMEWEEREIDN 621
Qy 602 SEIYTLIEESONQOEKNEOELBELDKWASLMMFDTITNMLWY 643
Db 622 SIYSLIEESONQOEKNEOELBELDKWASLMMFDTITNMLWY 663

```

RESULT 10

```

ID 056110 PRELIMINARY; PRT; 851 AA.
AC 056110;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_Taxid=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=SEWHS.1;
RX MEDLINE=98178716; PubMed=9519894;
RA McCutchan F.E., Sanders-Buell E., Salminen M.O., Carr J.K.,
   Sheppard W.H.;
   "Diversity of the human immunodeficiency virus type 1 envelope
   glycoprotein in San Francisco Men's Health Study participants.";
   AIDS Res. Hum. Retroviruses 14:329-337(1998).
DR EMBL: AF025756; AAC40593.1; -.
DR InterPro: IPR000328; Env GP41.
DR InterPro: IPR00173; GAP dhodrogenase.
DR InterPro: IPR00777; GP120.
DR Pfam: PF00517; GP120; 1.
DR Pfam: PF00517; GP41; 1.
DR PROSITE: PS00071; GAPDH; 1.
DR AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 851 AA; 96516 MW; 1CDAC83EE6464531 CRC64;

```

Query Match 88.4%; Score 3059; DB 15; Length 851;

Best Local Similarity 88.9%; Pred. No. 7,2e-241; Indels 8; Gaps 4;

```

Matches 576; Conservative 28; Mismatches 36; Indels 8; Gaps 4;
Qy 2 EKLWTVYVYGVVWKEATTTLFCASDAKAYDEVHNVATHACVPTDPNPOEVLLENVTE 61
Db 31 EKLWTVYVYGVVWKEATTTLFCASDAKAYDEVHNVATHACVPTDPNPOEVLLENVTE 90
Qy 62 HFNMMKNMVEOMQEDIIISLDOSLKPCKVLTPLCVTLNCKDVNATNTTNDSEGMERGE 120

```

```

Db      91  NFNMMKNNMVEOMEDIISLDOSLKPCKLTPLCVTLANCIDLRDNTTSSGGETMERG 150
Qy      121  EIKNCSFNITTSIRDEVOKEYALFYKLDVVPIDNNNTSYRLISCDTSVITQACPKISFEP 180
Db      151  EIKNCSFNITTSIRDKVQKEYALLHKLDPVID-NDTSRLVSCNITSVITQACPKISFEP 209
Qy      181  IPIHCAPAGPAILKCNCKTFNGKPCKNVSTVOCTHGIRPVSTOLLNGLSABEYVI 240
Db      210  IPIHCAPAGPAILKCNCKTFNGKPCKNVSTVOCTHGIRPVSTOLLNGLSABEYVI 269
Qy      241  RSDNFTNNAKTIIVOLKESVEINCTRPNNTKSHIGBRAFYTTGELIIGIRQACHNI 300
Db      270  RSDNFTNNAKTIIVOLKESVEINCTRPNNTKSHIGBRAFYTTGELIIGIRQACHNI 329
Qy      301  SPAKNDTLKOIVIKLREOFENKTIYFNHSSGDEPEIVMHSFNCGEFFYCNSTOLFENST 360
Db      330  SPAKNDTLKOIVIKLREOFENKTIYFNHSSGDEPEIVMHSFNCGEFFYCNSTOLFENST 389
Qy      361  WN-----NNTGSSNTEGNTITLPCRIKQIIMMOEGKAMVAPPIRQIRCSSNITGL 415
Db      390  WSNSTMTNDTEGVNNT-GENITLPCRIKQIIMMOEGKAMVAPPIRQIRCSSNITGL 448
Qy      416  LTRDGINENGTETIRPFGGDMRDMRSFYYKYKVIKIEPLGVAIPKCKRRVVOBKAV 475
Db      449  LTRDGINENGTETIRPFGGDMRDMRSFYYKYKVIKIEPLGVAIPKCKRRVVOBKAV 508
Qy      476  GIGAFGLGFLGAAGSTMGAASMTLTVQARLLSGIVQOONNLRAIEAQORLOLTWMI 535
Db      509  GIGAFGLGFLGAAGSTMGAASMTLTVQARLLSGIVQOONNLRAIEAQORLOLTWMI 568
Qy      536  KOLQARVLAVERYLDQOQLLGIMGCSGKLICTTAVPMNWSNKSIDRIIMNMWEMER 595
Db      569  KOLQARVLAVERYLDQOQLLGIMGCSGKLICTTAVPMNWSNKSIDRIIMNMWEMER 628
Qy      596  EIDNTSEYTLIESONOEKNEOELLELDKASLMMNFDITNMLY 643
Db      629  EINNTSLIYTLIESONOEKNEOELLELDKASLMMNFDITNMLY 676

```

RESULT 11

```

ID      077989  PRELIMINARY; PRT; 863 AA.
AC      077989;
DT      01-NOV-1996 (TREMBLrel. 01, Created)
DT      01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT      01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE      Envelope glycoprotein.
GN      ENV.
OS      Human immunodeficiency virus 1.
OC      Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX      NCBI_TaxID=11676;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=SF128A;
RX      MEDLINE=91056585; PubMed=2243391;
RA      Liu Z.-Q., Wood C., Levy J.A., Cheng-Mayer C.;
RT      "The viral envelope gene is involved in macrophage tropism of a human
RT      immunodeficiency virus type 1 strain isolated from brain tissue.";
RL      J. Virol. 64:6148-6153(1990).
DR      EMBL; M95292; AAA44331.1; -
DR      InterPro; IPR000328; Env_Gp1.
DR      Pfam; PF00516; GP120.1.
DR      Pfam; PF00517; GP41.1.
KW      AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ      SEQUENCE 863 AA; 98222 MW; 0157397FAB84C5D7 CRC64;

```

```

Query Match      88.0%; Score 3048; DB 15; Length 863;
Best Local Similarity 87.0%; Pred. No. 5.8e-240;
Matches 574; Conservative 26; Mismatches 40; Indels 20; Gaps 5;

```

```

2 EKLAATVYGVPMKEATTTLCASDAKAYDTEVHNWATACVPTDPNPGSEVLENVTE 61
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

Db      31  EKLAATVYGVPMKEATTTLCASDAKAYDTEVHNWATACVPTDPNPGSEVLENVTE 90
Qy      62  HFNMMKNNMVEOMEDIISLDOSLKPCKLTPLCVTLANCKD-----VNAATNTNDEGTM 117
Db      91  NFNMMKNNMVEOMEDIISLDOSLKPCKLTPLCVTLANCKDNDLRDNT-TNDTNATIS 149
Qy      118  -----ERGEIKNCSFNITTSIRDEVOKEYALFYKLDVVPIDNNNTSYRLISCDTSVITQA 172
Db      150  SLSRGETGEIKNCSFNITTSIRDKVQKEYALFYKLDVVPIDNNNTSYRLINCNSTITQA 209
Qy      173  CPKISFEPPIHYCAPAGPAILKCNCKTFNGKPCKNVSTVOCTHGIRPVSTOLLNGLS 232
Db      210  CPKISFEPPIHYCAPAGPAILKCNCKTFNGKPCKNVSTVOCTHGIRPVSTOLLNGLS 269
Qy      233  LAEEVVRSDNFTNNAKTIIVOLKESVEINCTRPNNTKSHIGBRAFYTTGELIIGD 292
Db      270  LAEEVVRSDNFTNNAKTIIVOLKESVEINCTRPNNTKSHIGBRAFYTTGELIIGD 329
Qy      293  IROACHNISRAKMDTLKOIVIKLREOFENKTIYFNHSSGDEPEIVMHSFNCGEFFYCN 352
Db      330  IROACHNISRAKMDTLKOIVIKLREOFENKTIYFNHSSGDEPEIVMHSFNCGEFFYCN 389
Qy      353  STOLFENSTWN-----NNTGSSNTEGNTITLPCRIKQIIMMOEGKAMVAPPIRG 403
Db      390  STOLFENSTWN-----NNTGSSNTEGNTITLPCRIKQIIMMOEGKAMVAPPIRG 448
Qy      404  QIRCSNITGLLTRDGINENGTETIRPFGGDMRDMRSFYYKYKVIKIEPLGVAIPK 463
Db      449  QIRCSNITGLLTRDGINENGTETIRPFGGDMRDMRSFYYKYKVIKIEPLGVAIPK 508
Qy      464  KRRVVOBKRAVIGAVLGLGFLGAAGSTMGAASMTLTVQARLLSGIVQOONNLRAIEA 523
Db      509  KRRVVOBKRAVIGAVLGLGFLGAAGSTMGAASMTLTVQARLLSGIVQOONNLRAIEA 568
Qy      524  QORMLQLTWMIKOLQARVLAVERYLDQOQLLGIMGCSGKLICTTAVPMNWSNKSIDR 583
Db      569  QORMLQLTWMIKOLQARVLAVERYLDQOQLLGIMGCSGKLICTTAVPMNWSNKSIDR 628
Qy      584  IMNMWEMEREIDNTSEYTLIESONOEKNEOELLELDKASLMMNFDITNMLY 643
Db      629  IMNMWEMEREIDNTSEYTLIESONOEKNEOELLELDKASLMMNFDITNMLY 688

```

RESULT 12

```

ID      09WJMS  PRELIMINARY; PRT; 868 AA.
AC      09WJMS;
DT      01-NOV-1999 (TREMBLrel. 12, Created)
DT      01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT      01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE      Envelope polyprotein.
GN      ENV.
OS      Human immunodeficiency virus 1.
OC      Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX      NCBI_TaxID=11676;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      Pang G., Weiser B., Chappey C., Visokey A., Townsend L., Wang Q.,
RA      Burger H.;
RT      "Complete Plasma HIV-1 Sequence: Sp1-Promoter Deletion Can lead to
RT      Non-Progressive Infection.";
RL      Submitted (SEP-1996) to the EMBL/Genbank/DBJ databases.
DR      EMBL; U69584; AAD10875.1; -
DR      InterPro; IPR000328; Env_Gp1.
DR      InterPro; IPR000777; GP120.
DR      Pfam; PF00516; GP120.1.
DR      Pfam; PF00517; GP41.1.
KW      AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ      SEQUENCE 868 AA; 98569 MW; F2D138F08BEFC02 CRC64;

```

```

Query Match      88.0%; Score 3045.5; DB 15; Length 868;
Best Local Similarity 87.7%; Pred. No. 9.4e-240;
Matches 582; Conservative 22; Mismatches 37; Indels 23; Gaps 7;

```

```

QY 2 EKLWTVYGVVWPKKATTTLLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTTE 61
D 31 EQLWTVYGVVWPKKATTTLLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTTE 90
QY 62 HFNWKNWVQWQEDDIISLMDOSLKPCVKLTPLCTVLNCKD-----VNATNTDSE 114
D 91 NFNWKNWVQWQEDDIISLMDOSLKPCVKLTPLCTVLNCKDKNATVKNATVNNSSW 150
QY 115 GTRMERGEIKNCSFNITTSIRDEVOKEVALFYKLDVPIDN-----NNTSYLLISDGT 166
D 151 GGMERGEIKNCSFNITTSIRDEVOKEVALFYKLDVPIDNADNNNTTNTSYLLISDGT 210
QY 167 SVITQACPKISFEPPIHYCAPAGFAILKNDKTFNGKGPCKNVSTVOCTGIRPVVSTQ 226
D 211 SVITQACPKISFEPPIHYCAPAGFAILKNDKTFNGKGPCKNVSTVOCTGIRPVVSTQ 270
QY 227 LLLNGSLAEEVVRISDNFTNNAKTIIVOLKESVEINCTRPNNTRKSIHIGPGRAYTT 286
D 271 LLLNGSLAEEVVRISDNFTNNAKTIIVOLKESVEINCTRPNNTRKSIHIGPGRAYTT 330
QY 287 GEIGDIRQAHCTNKRTEMNNTLAKITEKREQFNNTTIVFNHSSGGDEPEIVHSPICG 390
D 331 GEIGDIRQAHCTNKRTEMNNTLAKITEKREQFNNTTIVFNHSSGGDEPEIVHSPICG 390
QY 346 GEFYCNSTOLFNSTWN---NNTGSSNTEGNTITLPCRIOIIMWQEVGKAMVAPPIR 402
D 391 GEFYCNSTOLFNSTWNSTGNISSDTERRN-TTLPCRIOIIMWQEVGKAMVAPPIR 449
QY 403 GOIRCSSNITGLLTRDGIN--ENGTEIPRPGGDMNDNRSEFYKVKVIEPLGVAP 460
D 450 GOIRCSSNITGLLTRDGIN--ENGTEIPRPGGDMNDNRSEFYKVKVIEPLGVAP 509
QY 461 TKCRVRVQREKRVG-IGAVPLGFLGAAGSTMGASMTLVQARLLSGIVQOONNLLR 519
D 510 TKARVRVQREKRVG-IGAVPLGFLGAAGSTMGASMTLVQARLLSGIVQOONNLLR 569
QY 520 AIEAQRMLQTLVWGIKOLQARVLAVERLYLGDQOLGIMGCSGKLICTTAVPWNASMKN 579
D 570 AIEAQRMLQTLVWGIKOLQARVLAVERLYLGDQOLGIMGCSGKLICTTAVPWNASMKN 629
QY 580 SLDRIMNNTWMEWEREIDNTSEIYLLIESONOEKNEBELLDKVASLNNPDIITN 639
D 630 SLDRIMNNTWMEWEREIDNTSEIYLLIESONOEKNEBELLDKVASLNNPDIITN 689
QY 640 WLMY 643
D 690 WLMY 693

```

RESULT 13

092762 PRELIMINARY; PRT; 854 AA.

AC 092762; 01-NOV-1998 (Tremblrel. 08, Created)

DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)

DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)

DE Envelope glycoprotein.

GN ENV.

OS Human immunodeficiency virus 1.

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OX NCBI_TaxID=11676; [1]

RN SEQUENCE FROM N.A.

RA STRAIN=SEWHS.3; MEDLINE=98178716; PubMed=9519894; McCutchan F.E.; Sanders-Buell E.; Salminen M.O.; Carr J.K.; Sheppard W.H.;

RT "diversity of the human immunodeficiency virus type 1 envelope glycoprotein in San Francisco Men's Health Study participants";

RL AIDS Res. Hum. Retroviruses 14:329-337(1998).

DR EMBL; AF025754; AAC40591.1; InterPro; IPR000328; Env_GP41.

```

DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polypeptide; Transmembrane.
SQ SEQUENCE 854 AA; 97287 MW; AFD8A82B62C8B2D CRC64;

Query Match      87.9%; Score 3044; DB 15; Length 854;
Best Local Similarity 88.3%; Pred. No. 1,2e-239;
Matches 573; Conservative 29; Mismatches 39; Indels 8; Gaps 5;

QY 2 EKLWTVYGVVWPKKATTTLLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTTE 61
D 31 EQLWTVYGVVWPKKATTTLLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTTE 91
QY 62 HFNWKNWVQWQEDDIISLMDOSLKPCVKLTPLCTVLNCKD-----VNATNTDSE 119
D 91 NFNWKNWVQWQEDDIISLMDOSLKPCVKLTPLCTVLNCKDKNATVKNATVNNSSW 151
QY 120 GEIKNCSFNITTSIRDEVOKEVALFYKLDVPIDNNTSYRLISDTSVITQACPKISFE 179
D 152 GEIKNCSFNITTSIRDEVOKEVALFYKLDVPIDNNTSYRLISDTSVITQACPKISFE 211
QY 180 PIPHYCAPAGFAILKNDKTFNGKGPCKNVSTVOCTGIRPVVSTOLLNGSLAEEV 239
D 212 PIPHYCAPAGFAILKNDKTFNGKGPCKNVSTVOCTGIRPVVSTOLLNGSLAEEV 271
QY 240 IRSNFTNNAKTIIVOLKESVEINCTRPNNTRKSIHIGPGRAYTTGEIGDIRQAH 299
D 272 IRSNFTNNAKTIIVOLKESVEINCTRPNNTRKSIHIGPGRAYTTGEIGDIRQAH 331
QY 300 ISRAKNDTLQOIVIKLREOFENKTIIVFNHSSGGDEPEIVHSPICGGEFFYCNSTOLFNS 359
D 332 ISRAKNDTLQOIVIKLREOFENKTIIVFNHSSGGDEPEIVHSPICGGEFFYCNSTOLFNS 391
QY 360 TWNN-NTEGSSNTEGNTITLPCRIOIIMWQEVGKAMVAPPIRGOIRCSSNITGLLT 417
D 392 TWNN-NTEGSSNTEGNTITLPCRIOIIMWQEVGKAMVAPPIRGOIRCSSNITGLLT 451
QY 418 RDGINENGT---EIPRPGGDMNDNRSEFYKVKVIEPLGVAPFKCRVRVQREKRA 474
D 452 RDGINENGT---EIPRPGGDMNDNRSEFYKVKVIEPLGVAPFKCRVRVQREKRA 510
QY 475 VGIGAVPLGFLGAAGSTMGASMTLVQARLLSGIVQOONNLLRAIEAQRMLQTLVW 534
D 511 VGIGAVPLGFLGAAGSTMGASMTLVQARLLSGIVQOONNLLRAIEAQRMLQTLVW 570
QY 535 IKOLQARVLAVERLYLGDQOLGIMGCSGKLICTTAVPWNASMKNSLDRIMNNTWME 594
D 571 IKOLQARVLAVERLYLGDQOLGIMGCSGKLICTTAVPWNASMKNSLDRIMNNTWME 630
QY 595 REIDNTSEIYLLIESONOEKNEBELLDKVASLNNPDIITN 643
D 631 REIDNTSEIYLLIESONOEKNEBELLDKVASLNNPDIITN 679

```

RESULT 14

09YPS0 PRELIMINARY; PRT; 860 AA.

AC 09YPS0; 01-MAY-1999 (Tremblrel. 10, Created)

DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)

DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)

DE Envelope glycoprotein.

GN ENV.

OS Human immunodeficiency virus 1.

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OX NCBI_TaxID=11676; [1]

RN SEQUENCE FROM N.A.

RA Fang G.; Weiser B.; Chappey C.; Visovsky A.; Townsend L.; Wang Q.; Burger H.;

RT "Complete Plasma HIV-1 Sequence: Spl-Promoter Deletion Can Lead to Non-Progressive Infection.";

RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL: U69585; AAD10883.1; -

DR InterPro: IPR000328; Env GP1.

DR InterPro: IPR000777; GP120.

DR Pfam: PF00516; GP120; 1.

DR Pfam: PF00517; GP1; 1.

DR K1D5; Coat protein; Glycoprotein; Polyprotein; Transmembrane.

KM AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.

SEQUENCE 860 AA; 97662 MW; A5AC60C558B90BC CRC64;

Query Match 87.8%; Score 3039.5; DB 15; Length 860;

Best Local Similarity 87.8%; Pred. No. 2.9e-239;

Matches 577; Conservative 24; Mismatches 39; Indels 17; Gaps 5;

```

QY 2 EKLAVTVYGVVPMKEATTTLFCASDAKAYDTEVHNVNATHACVPTDNPQEVLENTYE 61
DB 31 EQLAVTVYGVVPMKEATTTLFCASDAKAYDTEVHNVNATHACVPTDNPQEVLENTYE 90
QY 62 HNNMKNMNMVEQMHEDIISLMDQSLKPCVKLTPLCVTLNCKD-----VNATNTNSE 114
DB 91 HNNMKNMNMVEQMHEDIISLMDQSLKPCVKLTPLCVTLNCKD-----VNATNTNSE 150
QY 115 GEMERGEIKNCSFNITTSIRDVOKEVALFYKLDVVPIDN-----NNTSYRLISCDTSYI 169
DB 151 EKMERGEIKNCSFNITTSIRDVOKEVALFYKLDVVPIDN-----NNTSYRLISCDTSYI 210
QY 170 TQACPKISFPIPIHYCAPAGFAILKCNKDTFNGKPCCKNVSTVOCTHGIRPVVSTQLL 229
DB 211 TQACPKISFPIPIHYCAPAGFAILKCNKDTFNGKPCCKNVSTVOCTHGIRPVVSTQLL 270
QY 220 NSGLAEVEVIRSDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPRAFYTTGEI 289
DB 271 NSGLAEVEVIRSDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPRAFYTTGEI 330
QY 290 IGDIOACHNISRANKNDTLKQIVIKLREOFENKTIIVFNHSSGSDPEIVNHSFCEGEF 349
DB 331 IGDIOACHNISRANKNDTLKQIVIKLREOFENKTIIVFNHSSGSDPEIVNHSFCEGEF 390
QY 350 YCNSTOLF--NSTNNNTSGNTEGNTITLPCRKOIINMVEGKAMYPPIRGQIR 407
DB 391 YCNSTOLF--NSTNNNTSGNTEGNTITLPCRKOIINMVEGKAMYPPIRGQIR 448
QY 408 SSNITGLLTRDGINENGTETFRPGGDMRDNRSEFYKVKYKIEPLGVAPTKCKRRV 467
DB 449 SSNITGLLTRDGINENGTETFRPGGDMRDNRSEFYKVKYKIEPLGVAPTKCKRRV 508
QY 468 VQREKRAVG-IGAVFLGFLGAGSTMGASMTLTVOARLLSGIVQOQNNLLRAIEAQO 526
DB 509 VQREKRAVG-IGAVFLGFLGAGSTMGASMTLTVOARLLSGIVQOQNNLLRAIEAQO 568
QY 527 MLQLTWVGIKQOLQARVLAVERYLADQQLLGIGGCGKLICTTAVPNNASMSKSLDRIM 586
DB 569 MLQLTWVGIKQOLQARVLAVERYLADQQLLGIGGCGKLICTTAVPNNASMSKSLDRIM 628
QY 567 NMTMMEWEREINDYTSIYTLIESONQOEKNOELLELDKVASLMMNFDITNMLWY 643
DB 629 NMTMMEWEREINDYTSIYTLIESONQOEKNOELLELDKVASLMMNFDITNMLWY 685

```

RESULT 15

ID Q9WJVS PRELIMINARY; PRT; 861 AA.

AC Q9WJVS; 01-NOV-1999 (TREMBLrel. 12, Created)

DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)

DE 01-OCT-2002 (TREMBLrel. 22, Last annotation update)

ENV. Envelope polyprotein.

GN Human immunodeficiency virus 1.

OS Virus; Retrovirus; Retroviridae; Lentivirus.

OC NCBI_Taxid=11676;

OX [1]

RN [1]

RP SEQUENCE FROM N.A.

FA Fang G., Weiser B., Chappey C., Visosky A., Townsend L., Wang Q.,

RA Burger H.; "Complete Plasma HIV-1 Sequence: Spl-Promoter Deletion Can Lead to

RT Non-Progressive Infection."; Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL: U69587; AAD10889.1; -

DR InterPro: IPR000328; Env GP1.

DR InterPro: IPR000777; GP120.

DR Pfam: PF00516; GP120; 1.

DR Pfam: PF00517; GP1; 1.

DR K1D5; Coat protein; Glycoprotein; Polyprotein; Transmembrane.

KM AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.

SEQUENCE 861 AA; 97993 MW; 3E3EFC977CC18EA CRC64;

Query Match 87.8%; Score 3039; DB 15; Length 861;

Best Local Similarity 87.7%; Pred. No. 3.1e-239;

Matches 577; Conservative 25; Mismatches 38; Indels 18; Gaps 6;

```

QY 2 EKLAVTVYGVVPMKEATTTLFCASDAKAYDTEVHNVNATHACVPTDNPQEVLENTYE 61
DB 31 EQLAVTVYGVVPMKEATTTLFCASDAKAYDTEVHNVNATHACVPTDNPQEVLENTYE 90
QY 62 HNNMKNMNMVEQMHEDIISLMDQSLKPCVKLTPLCVTLNCKD-----VNATNTNSE 114
DB 91 HNNMKNMNMVEQMHEDIISLMDQSLKPCVKLTPLCVTLNCKD-----VNATNTNSE 150
QY 115 GEMERGEIKNCSFNITTSIRDVOKEVALFYKLDVVPIDN-----NNTSYRLISCDTSYI 168
DB 151 EKMERGEIKNCSFNITTSIRDVOKEVALFYKLDVVPIDN-----NNTSYRLISCDTSYI 210
QY 169 ITQACPKISFPIPIHYCAPAGFAILKCNKDTFNGKPCCKNVSTVOCTHGIRPVVSTQLL 228
DB 211 ITQACPKISFPIPIHYCAPAGFAILKCNKDTFNGKPCCKNVSTVOCTHGIRPVVSTQLL 270
QY 229 NSGLAEVEVIRSDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPRAFYTTGEI 288
DB 271 NSGLAEVEVIRSDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPRAFYTTGEI 330
QY 289 IGDIOACHNISRANKNDTLKQIVIKLREOFENKTIIVFNHSSGSDPEIVNHSFCEGEF 348
DB 331 IGDIOACHNISRANKNDTLKQIVIKLREOFENKTIIVFNHSSGSDPEIVNHSFCEGEF 390
QY 349 FYCNSTOLF--NSTNNNTSGNTEGNTITLPCRKOIINMVEGKAMYPPIRGQIR 406
DB 391 FYCNSTOLF--NSTNNNTSGNTEGNTITLPCRKOIINMVEGKAMYPPIRGQIR 448
QY 407 SSNITGLLTRDGINENGTETFRPGGDMRDNRSEFYKVKYKIEPLGVAPTKCKRRV 466
DB 449 SSNITGLLTRDGINENGTETFRPGGDMRDNRSEFYKVKYKIEPLGVAPTKCKRRV 508
QY 467 VQREKRAVG-IGAVFLGFLGAGSTMGASMTLTVOARLLSGIVQOQNNLLRAIEAQO 525
DB 509 VQREKRAVG-IGAVFLGFLGAGSTMGASMTLTVOARLLSGIVQOQNNLLRAIEAQO 568
QY 526 MLQLTWVGIKQOLQARVLAVERYLADQQLLGIGGCGKLICTTAVPNNASMSKSLDRIM 585
DB 569 MLQLTWVGIKQOLQARVLAVERYLADQQLLGIGGCGKLICTTAVPNNASMSKSLDRIM 628
QY 586 NMTMMEWEREINDYTSIYTLIESONQOEKNOELLELDKVASLMMNFDITNMLWY 643
DB 629 NMTMMEWEREINDYTSIYTLIESONQOEKNOELLELDKVASLMMNFDITNMLWY 686

```

Search completed: December 12, 2003, 12:32:36

Job time : 36.0726 secs

THIS PAGE BLANK (US 710)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: December 12, 2003, 12:24:08 ; Search time 34.4829 Seconds
(without alignments)
2665.163 Million cell updates/sec

Title: US-10-032-162-15

Perfect score: 3129
Sequence: 1 VEKLMVTYYGVVPWKKEAT.....ELDKKASIMNPFDTNMLMY 579

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

1 number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A.GeneSeq_19Jun03:*

- 1: /SIDS1/gcgcdata/geneSeq/geneSeq-emb1/AA1980.DAT:*
- 2: /SIDS1/gcgcdata/geneSeq/geneSeq-emb1/AA1981.DAT:*
- 3: /SIDS1/gcgcdata/geneSeq/geneSeq-emb1/AA1982.DAT:*
- 4: /SIDS1/gcgcdata/geneSeq/geneSeq-emb1/AA1983.DAT:*
- 5: /SIDS1/gcgcdata/geneSeq/geneSeq-emb1/AA1984.DAT:*
- 6: /SIDS1/gcgcdata/geneSeq/geneSeq-emb1/AA1985.DAT:*
- 7: /SIDS1/gcgcdata/geneSeq/geneSeq-emb1/AA1986.DAT:*
- 8: /SIDS1/gcgcdata/geneSeq/geneSeq-emb1/AA1987.DAT:*
- 9: /SIDS1/gcgcdata/geneSeq/geneSeq-emb1/AA1988.DAT:*
- 10: /SIDS1/gcgcdata/geneSeq/geneSeq-emb1/AA1989.DAT:*
- 11: /SIDS1/gcgcdata/geneSeq/geneSeq-emb1/AA1990.DAT:*
- 12: /SIDS1/gcgcdata/geneSeq/geneSeq-emb1/AA1991.DAT:*
- 13: /SIDS1/gcgcdata/geneSeq/geneSeq-emb1/AA1992.DAT:*
- 14: /SIDS1/gcgcdata/geneSeq/geneSeq-emb1/AA1993.DAT:*
- 15: /SIDS1/gcgcdata/geneSeq/geneSeq-emb1/AA1994.DAT:*
- 16: /SIDS1/gcgcdata/geneSeq/geneSeq-emb1/AA1995.DAT:*
- 17: /SIDS1/gcgcdata/geneSeq/geneSeq-emb1/AA1996.DAT:*
- 18: /SIDS1/gcgcdata/geneSeq/geneSeq-emb1/AA1997.DAT:*
- 19: /SIDS1/gcgcdata/geneSeq/geneSeq-emb1/AA1998.DAT:*
- 20: /SIDS1/gcgcdata/geneSeq/geneSeq-emb1/AA1999.DAT:*
- 21: /SIDS1/gcgcdata/geneSeq/geneSeq-emb1/AA2000.DAT:*
- 22: /SIDS1/gcgcdata/geneSeq/geneSeq-emb1/AA2001.DAT:*
- 23: /SIDS1/gcgcdata/geneSeq/geneSeq-emb1/AA2002.DAT:*
- 24: /SIDS1/gcgcdata/geneSeq/geneSeq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3129	100.0	579	22	AAB61506
2	3056	97.7	643	22	AAB61505
3	2923	93.4	625	22	AAB61507
4	2800.5	89.5	855	20	AAW1581
5	2800.5	89.5	855	20	AAW88113
6	2782	88.9	883	22	AAB82761
7	2776.5	88.7	850	16	AAK67724
8	2771	88.6	619	23	AAU75156
9			646	23	AAU75155

10	2764	88.3	847	21	AAV97073
11	2759	88.2	842	24	ABU66565
12	2749	87.9	842	23	ABU66211
13	2735.5	87.4	851	9	AAP80967
14	2725	87.1	857	16	AAB67725
15	2721.5	87.0	868	7	AAP60422
16	2719	86.9	856	22	AAB85999
17	2718.5	86.9	863	14	AAR43869
18	2718.5	86.9	868	23	AAO1389
19	2717.5	86.8	868	7	AAP60063
20	2715	86.8	856	7	AAP61514
21	2715	86.8	856	20	AAW89325
22	2714	86.7	855	19	AAW53112
23	2714	86.7	855	21	AAV77298
24	2714	86.7	855	21	AAV77302
25	2714	86.7	855	24	ABU57550
26	2714	86.7	855	24	ABU57553
27	2714	86.7	863	7	AAP61509
28	2714	86.7	863	13	AAR29706
29	2713.5	86.7	855	12	AAR14905
30	2713	86.7	856	7	AAP60131
31	2712	86.7	854	21	AAB10697
32	2712	86.7	854	21	AAB10053
33	2712	86.7	854	22	AAB86199
34	2712	86.7	854	22	AAB67277
35	2712	86.7	854	23	ABG73663
36	2711	86.6	856	13	AAR25940
37	2709	86.6	863	7	AAP60349
38	2708.5	86.6	860	18	AAW31284
39	2707.5	86.5	880	18	AAW23333
40	2707.5	86.5	880	20	AAW73332
41	2707.5	86.5	880	21	AAW14844
42	2706.5	86.5	856	16	AAB67726
43	2706.5	86.5	855	19	AAW43069
44	2704.5	86.4	901	8	AAP70665
45	2702	86.4	856	21	AAV97072

ALIGNMENTS

RESULT 1	
AAB61506	
ID	AAB61506 standard; Protein; 579 AA.
AC	AAB61506;
DT	05-APR-2001 (first entry)
DE	HIV-1 delcav12* SOS gp140 glycoprotein.
KM	gp120; anti-HIV; gene therapy; vaccine; viral envelope protein; gp140.
KW	gp140.
OS	Human immunodeficiency virus type 1.
PN	WO200100648-A1.
PD	04-JAN-2001.
PF	23-JUN-2000; 2000WO-US17267.
PR	25-JUN-1999; 99US-0340992.
PA	(PROG-) PROGENICS PHARM INC.
PA	(AARO-) AARON DIAMOND AIDS RES CENT.
PI	Binley JM, Schuelke N, Olson WC, Madden PJ, Moore JP;
DR	WPI; 2001-122993/13.
XX	N-PSDB; AAF28582.
PT	New viral envelope proteins, useful for producing vaccines to treat

Variant HIV-1 SF16
Human immunodeficiency virus isolate SF
HIV Env isolate SF
HIV protein HT6.
gp120 from the HIV
Sequence of LAV V1
Amino acid sequenc
HIV-III ENV-LOR g
Lymphadenopathy-as
HIV virus env gene
Sequence of enviro
HIV-1 env protein
ENV protein contai
HIV-1 (ATCC CRL 85
HIV-1 (ATCC CRL 85
AIDS associated re
Sequence of ARV-2
env gene decoded f
HIV-1 BA-L clone-e
Sequence of the AI
HIV-1 env protein.
HIV-1 coat protein
HIV gp41 DNA SEQ I
Protein encoded by
HIV-1 NL4-3 gp120
Modified HIV env g
HIV-III virus (HI
HIV-SF2 virus gp12
Human immunodefici
HIV envelope prote
gp120 from the HIV
HIV-1 gp120 protei
Sequence encoded b
Wild type HIV-1 HX

PT human immunodeficiency virus-1 infections, comprises amino acid
PT sequence mutations such that viral transmembrane-surface protein
PT complex is more stable -

PS Disclosure; Fig 14; 109pp; English.

CC The present invention relates to a viral envelope protein. The viral
CC envelope protein comprises a viral surface protein (e.g. glycoprotein
CC gp120) and a corresponding viral transmembrane protein (e.g. gp41), in
CC which the viral envelope protein contains one or more amino acid sequence
CC mutations that enhance the stability of the complex formed between the
CC viral surface and transmembrane proteins. The viral envelope protein can
CC be used in the treatment of viral infection e.g. HIV-1 infection. The
CC present invention is HIV-1 deltaV1V2* SOS gp140, which was used in the
CC present invention.

SO Sequence 579 AA;

Query Match 100.0%; Score 3129; DB 22; Length 579;
Best Local Similarity 100.0%; Pred. No. 6,4e-146;
Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEKLMVTVYGVVPWKATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENT 60
DB 1 VEKLMVTVYGVVPWKATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENT 60
QY 61 EHFMMKNNWVEQWQEDIIISLMDOSLKPCKVLTPLCGAGCDSVTQACPKISEPPIH 120
DB 61 EHFMMKNNWVEQWQEDIIISLMDOSLKPCKVLTPLCGAGCDSVTQACPKISEPPIH 120
QY 121 YCABAGFALIKCNKDTFNGKGPCKNVSTVQCTHGIRPVVSTOLLNGSLAESEVIRSDN 180
DB 121 YCABAGFALIKCNKDTFNGKGPCKNVSTVQCTHGIRPVVSTOLLNGSLAESEVIRSDN 180
QY 181 FTNNAKTIIVOLAKSEVEINCTRPNNNTKSHIGRGAFTYTGIGIRQAHCHISRAK 240
DB 181 FTNNAKTIIVOLAKSEVEINCTRPNNNTKSHIGRGAFTYTGIGIRQAHCHISRAK 240
QY 241 WNDTLKOIVIKLRQEFENKTIIVFNHSSGGDEIIVHSSFGCEFFYCNSTOLFENSTWNN 300
DB 241 WNDTLKOIVIKLRQEFENKTIIVFNHSSGGDEIIVHSSFGCEFFYCNSTOLFENSTWNN 300
QY 301 TEGSNTTEGNTITLPCRIRKOIINMWQEVGKAMVAPPIRGQIRCSNITGLLTRDGINE 360
DB 301 TEGSNTTEGNTITLPCRIRKOIINMWQEVGKAMVAPPIRGQIRCSNITGLLTRDGINE 360
QY 361 NGTEIFRPGGDMRDNWSELYKYKVVKIEPLGVAFTCKRRVVOREKRAVGIGAVFLGF 420
DB 361 NGTEIFRPGGDMRDNWSELYKYKVVKIEPLGVAFTCKRRVVOREKRAVGIGAVFLGF 420
QY 421 LGAAGSTWGAASMTLTVOARLLLSGIVQOONNLRAIEAQORMLQITWGIKQIARVLA 480
DB 421 LGAAGSTWGAASMTLTVOARLLLSGIVQOONNLRAIEAQORMLQITWGIKQIARVLA 480
QY 481 VERVLGDOQLGIVGCSGKLICTAVPNNASMSKSLDRINNNMTWMEEREINDYSEI 540
DB 481 VERVLGDOQLGIVGCSGKLICTAVPNNASMSKSLDRINNNMTWMEEREINDYSEI 540
QY 541 YTLIEESQNOQKOELELDKASLWMPDITWMLY 579
DB 541 YTLIEESQNOQKOELELDKASLWMPDITWMLY 579

RESULT 2

AAB61505 ID AAB61505 standard; Protein; 643 AA.

XX AAB61505;

XX 05-APR-2001 (first entry)

DE HIV-1 SOS gp140 glycoprotein.

KW gp120; anti-HIV; gene therapy; vaccine; viral envelope protein; gp41;
KW gp140.

OS Human immunodeficiency virus type 1.

PN WO200100648-A1.

PD 04-JAN-2001.

PE 23-JUN-2000; 2000MO-US17267.

PR 25-JUN-1999; 99US-0340992.

PA (PROG-) PROGENICS PHARM INC.

XX (AARO-) AARON DIAMOND AIDS RES CENT.

XX Binley JM, Schuelke N, Olson WC, Madden PJ, Moore JB;

DR WPI; 2001-122993/13.

DR N-PSDB; AAF28581.

PT New viral envelope proteins, useful for producing vaccines to treat

PT human immunodeficiency virus-1 infections, comprises amino acid

PT sequence mutations such that viral transmembrane-surface protein

PT complex is more stable -

PS Disclosure; Fig 13; 109pp; English.

CC The present invention relates to a viral envelope protein. The viral
CC envelope protein comprises a viral surface protein (e.g. glycoprotein
CC gp120) and a corresponding viral transmembrane protein (e.g. gp41), in
CC which the viral envelope protein contains one or more amino acid sequence
CC mutations that enhance the stability of the complex formed between the
CC viral surface and transmembrane proteins. The viral envelope protein can
CC be used in the treatment of viral infection e.g. HIV-1 infection. The
CC present invention is HIV-1 SOS gp140, which was used in the present
CC invention.

SO Sequence 643 AA;

Query Match 97.7%; Score 3056; DB 22; Length 643;
Best Local Similarity 99.3%; Pred. No. 2,6e-142;
Matches 574; Conservative 0; Mismatches 5; Indels 64; Gaps 1;

QY 1 VEKLMVTVYGVVPWKATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENT 60
DB 1 VEKLMVTVYGVVPWKATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENT 60
QY 61 EHFMMKNNWVEQWQEDIIISLMDOSLKPCKVLTPLCGA-----GCDTSVTQACPKISEP 98
DB 61 EHFMMKNNWVEQWQEDIIISLMDOSLKPCKVLTPLCTLACKOVNATNTTNDSEGTWERG 120
QY 99 -----GCDTSVTQACPKISEP 116
DB 121 EIKNCSFNITTSIRDEVOKEYALFYKLDVVPIDNNNTSYRLISCDTSVITQACPKISEP 180
QY 117 IPIHYCAPAGFALIKCNKDTFNGKGPCKNVSTVQCTHGIRPVVSTOLLNGSLAESEVIR 176
DB 117 IPIHYCAPAGFALIKCNKDTFNGKGPCKNVSTVQCTHGIRPVVSTOLLNGSLAESEVIR 176
QY 181 IPIHYCAPAGFALIKCNKDTFNGKGPCKNVSTVQCTHGIRPVVSTOLLNGSLAESEVIR 240
DB 181 IPIHYCAPAGFALIKCNKDTFNGKGPCKNVSTVQCTHGIRPVVSTOLLNGSLAESEVIR 240
QY 177 RSDNFTNNAKTIIVOLAKSEVEINCTRPNNNTKSHIGRGAFTYTGIGIRQAHCHISRAK 236
DB 177 RSDNFTNNAKTIIVOLAKSEVEINCTRPNNNTKSHIGRGAFTYTGIGIRQAHCHISRAK 236
QY 241 RSDNFTNNAKTIIVOLAKSEVEINCTRPNNNTKSHIGRGAFTYTGIGIRQAHCHISRAK 300
DB 241 RSDNFTNNAKTIIVOLAKSEVEINCTRPNNNTKSHIGRGAFTYTGIGIRQAHCHISRAK 300
QY 237 SRKAMNDTLKOIVIKLRQEFENKTIIVFNHSSGGDEIIVHSSFGCEFFYCNSTOLFENST 296
DB 237 SRKAMNDTLKOIVIKLRQEFENKTIIVFNHSSGGDEIIVHSSFGCEFFYCNSTOLFENST 296
QY 301 SRKAMNDTLKOIVIKLRQEFENKTIIVFNHSSGGDEIIVHSSFGCEFFYCNSTOLFENST 360
DB 301 SRKAMNDTLKOIVIKLRQEFENKTIIVFNHSSGGDEIIVHSSFGCEFFYCNSTOLFENST 360
QY 297 WNNNTTEGNTTEGNTITLPCRIRKOIINMWQEVGKAMVAPPIRGQIRCSNITGLLTRD 356
DB 297 WNNNTTEGNTTEGNTITLPCRIRKOIINMWQEVGKAMVAPPIRGQIRCSNITGLLTRD 356
QY 361 GINENGTEIFRPGGDMRDNWSELYKYKVVKIEPLGVAFTCKRRVVOREKRAVGIGAV 416
DB 361 GINENGTEIFRPGGDMRDNWSELYKYKVVKIEPLGVAFTCKRRVVOREKRAVGIGAV 416

```

Db      421 GINENGTEIFRPGGDMNDNRSEFEYKYKVIIEPLGVAPTKCKRRVQREKRAVGIGAV 480
Qy      417 FLGFLGAAGSTMGASMTLTVOARLLLSGIYQOQNNLLRAIEAQRMQLTVMGIKQLQA 476
Db      461 FLGFLGAAGSTMGASMTLTVOARLLLSGIYQOQNNLLRAIEAQRMQLTVMGIKQLQA 540
Qy      477 RVLAVERYLGDOQLLGWCGSGKLICTTAVPWNASMSNKSIDRIINNNTMMEEREIDNY 536
Db      541 RVLAVERYLGDOQLLGWCGSGKLICTTAVPWNASMSNKSIDRIINNNTMMEEREIDNY 600
Qy      537 TSEIYTLIESONOEKNEOELELDKMASLNNMFDITNMLMY 579
Db      601 TSEIYTLIESONOEKNEOELELDKMASLNNMFDITNMLMY 643

RESULT 3
AAB61507
ID      AAB61507 standard; Protein; 625 AA.
Qy      AAB61507;
Db      AAB61507;
Dt      05-APR-2001 (first entry)
Xx      HIV-1 deltaV3 SOS gp140 glycoprotein.
Xx      gp120; anti-HIV; gene therapy; vaccine; viral envelope protein; gp14;
Xx      gp140.
Xx      Human immunodeficiency virus type 1.
Xx      WO200100648-A1.
Xx      04-JAN-2001.
Xx      23-JUN-2000; 2000MO-US17267.
Xx      25-JUN-1999; 99US-0340992.
Xx      (PROG-) PROGENICS PHARM INC.
Xx      (AARO-) AARON DIAMOND AIDS RES CENT.
Xx      Binley JM, Schuelke N, Olson WC, Maddon PJ, Moore JP;
Xx      WPI; 2001-122993/13.
Xx      N-PSDB; AAF26583.
Xx      New viral envelope proteins, useful for producing vaccines to treat
Xx      human immunodeficiency virus-1 infections, comprises amino acid
Xx      sequence mutations such that viral transmembrane-surface protein
Xx      complex is more stable
Xx      Disclosure; Fig 15; 109pp; English.
Xx      The present invention relates to a viral envelope protein. The viral
Xx      envelope protein comprises a viral surface protein (e.g. glycoprotein
Xx      gp120) and a corresponding viral transmembrane protein (e.g. gp41) in
Xx      which the viral envelope protein contains one or more amino acid sequence
Xx      mutations that enhance the stability of the complex formed between the
Xx      viral surface and transmembrane proteins. The viral envelope protein can
Xx      be used in the treatment of viral infection e.g. HIV-1 infection. The
Xx      present invention is HIV-1 deltaV3 SOS gp140, which was used in the
Xx      sequence 625 AA;
Qy      Query Match 93.4%; Score 2923; DB 22; length 625;
Db      Best Local Similarity 86.0%; Pred. No. 8.2e-136;
        Matches 553; Conservative 0; Mismatches 8; Indels 82; Gaps 2;
Qy      1 VEKLVATYYVGVVWKEATITLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENT 60
        1 VEKLVATYYVGVVWKEATITLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENT 60

```

```

Qy      61 EHFNMKNMVEQKQEDIIISLMDQSLKPCVKLTPLCA----- 98
Db      61 EHFNMKNMVEQKQEDIIISLMDQSLKPCVKLTPLCA----- 98
Qy      99 -----GCDSVTIQAQPKISFEP 116
Db      121 EIKNCSFNITTSIRDEVOKEYALFYKLDVVAIDNNNTSYRLISCDSVTITQAQPKISFEP 180
Qy      117 IPIHCAAGAPAILKCNKRTNGKPCQVSTVQCTHRIIPVVSQLLNGSLAEVEVI 176
Db      161 IPIHCAAGAPAILKCNKRTNGKPCQVSTVQCTHRIIPVVSQLLNGSLAEVEVI 240
Qy      177 RSDNFTNAKTIIVQLKESVEINCTRPNNNTKRSIHIGPRAFYTTGELIGDIRQAHNI 236
Db      241 RSDNFTNAKTIIVQLKESVEINCTRPNNNG-----AGDIRQAHNI 282
Qy      237 SRAKNDTLKQIVIKLRQFENKTIIVFNHSSGQDEIYVHSGFECGSEFFYCNSDTQFNST 296
Db      283 SRAKNDTLKQIVIKLRQFENKTIIVFNHSSGQDEIYVHSGFECGSEFFYCNSDTQFNST 342
Qy      297 WNNTEGNNTEGNTITLPCRIRKQIINMVOYGRKMAPIRGOIRCSSNTGLLTBDG 356
Db      343 WNNTEGNNTEGNTITLPCRIRKQIINMVOYGRKMAPIRGOIRCSSNTGLLTBDG 402
Qy      357 GINENGTEIFRPGGDMNDNRSEFEYKYKVIIEPLGVAPTKCKRRVQREKRAVGIGAV 416
Db      403 GINENGTEIFRPGGDMNDNRSEFEYKYKVIIEPLGVAPTKCKRRVQREKRAVGIGAV 462
Qy      417 FLGFLGAAGSTMGASMTLTVOARLLLSGIYQOQNNLLRAIEAQRMQLTVMGIKQLQA 476
Db      463 FLGFLGAAGSTMGASMTLTVOARLLLSGIYQOQNNLLRAIEAQRMQLTVMGIKQLQA 522
Qy      477 RVLAVERYLGDOQLLGWCGSGKLICTTAVPWNASMSNKSIDRIINNNTMMEEREIDNY 536
Db      523 RVLAVERYLGDOQLLGWCGSGKLICTTAVPWNASMSNKSIDRIINNNTMMEEREIDNY 582
Qy      537 TSEIYTLIESONOEKNEOELELDKMASLNNMFDITNMLMY 579
Db      583 TSEIYTLIESONOEKNEOELELDKMASLNNMFDITNMLMY 625

RESULT 4
AAM11581
ID      AAM11581 standard; Protein; 855 AA.
Xx      AAM11581;
Xx      25-MAR-2003 (updated)
Xx      25-MAR-1997 (first entry)
Xx      Human Immunodeficiency Virus-1 strain BA-L envelope protein.
Xx      Acquired immune deficiency syndrome; AIDS; envelope protein;
Xx      env gene; vaccine.
Xx      Human immunodeficiency virus type 1 (strain BA-L).
Xx      US5576000-A.
Xx      19-NOV-1996.
Xx      15-FEB-1995; 95US-0368809.
Xx      17-OCT-1990; 90US-0599491.
Xx      25-FEB-1993; 93US-0022835.
Xx      15-FEB-1995; 95US-0368809.
Xx      (USSH ) US DEPT HEALTH & HUMAN SERVICES.
Xx      Franchini G, Gallo RC, Gartner S, Lori FC, Markham PD;
Xx      Popovic M, Reitz MS;

```

DR WPI: 1997-011206/01.
 DR N-PSDB; AAT58551.
 PT New isolated envelope protein of HIV-1 strain BA-L and recombinant
 PT equivalents - useful as immunogens for vaccines and antibody prodn.,
 PT typical of US clinical isolates
 PS Claim 1; Fig 9; 86pp; English.
 XX
 CC A HindIII fragment of unintegrated viral DNA representing the BA-L
 CC genome was cloned into lambda phage Charon 28 DNA from total DNA of
 CC peripheral blood lymphocytes infected with and producing HIV-1 (BA-L).
 CC A positive clone was selected by hybridisation using a HIV-1 env
 CC probe. This clone, designated BA-L1, was found to contain the
 CC entire gene for the envelope protein on a 2.8 kb HindIII-XbaI
 CC fragment and a 0.4 kb EcoRI-HindIII fragment. When cloned together
 CC these fragments comprise the env gene, as well as the coding regions
 CC for rev and the rev-responsive element of env, both necessary for
 CC efficient expression in eukaryotic cells. The claimed recombinant
 CC produced envelope protein can be used as an immunogen for raising
 CC antibodies against HIV.
 CC (Updated on 25-MAR-2003 to correct PF field.)
 XX
 SQ Sequence 855 AA;
 Query Match 89.5%; Score 2800.5; DB 18; Length 855;
 Best Local Similarity 81.2%; Pred. No. 1e-129;
 Matches 527; Conservative 20; Mismatches 31; Indels 71; Gaps 2;
 QY 2 EKLWVTVYGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTE 61
 DB 31 EKLWVTVYGVVWKEATTTLFCASDRKAYDTEVHNWATHACVPTDNPQEVLENTE 90
 QY 62 HFNWKNWVQWQEDIIISLDQSLKPCVKLTPLC----- 96
 DB 91 NFNWKNWVQWQEDIIISLDQSLKPCVKLTPLC----- 150
 QY 97 --GAG----- 110
 DB 151 MVGGGEMKNCSEFNTTNRGKVKQEVAFYKLDIAPIDNNSNNRRLISCTSVITQACP 210
 QY 111 KISFEPIPIHYCAPAGFAILKNDKTEGKPCKNVSTVQCTHGIRPVVSTQLLNGSLA 170
 DB 211 KVSFEPIPIHYCAPAGFAILKCKDKKFKGKPCSTVSTVQCTHGIRPVVSTQLLNGSLA 270
 QY 171 EEEVIRSDNFTNNAKTIIVQKESVEINCRPNNTKSIHIGGRAPYTTGELIGDIR 230
 DB 271 EEEVIRSDNFTNNAKTIIVQKESVEINCRPNNTKSIHIGGRAPYTTGELIGDIR 330
 QY 231 QAHGNIISPAKNDTLKQIVIKLREOFENKTIIVFNHSGGDEIIVHSGGGEFFYCNST 290
 DB 331 QAHGNIISPAKNDTLKQIVIKLREOFENKTIIVFNHSGGDEIIVHSGGGEFFYCNST 390
 QY 291 QLFNSTMNNTGSGNNTGNTTTLPCRIKQIIMNQEVGKAMYAPIRGOIRCSNITGL 350
 DB 391 QLFNSTMNNTGSGNNTGNTTTLPCRIKQIIMNQEVGKAMYAPIRGOIRCSNITGL 450
 QY 351 LITDDGGINENGTEIFRRPGGDMRNWSELYKYRVNKLIEPLGVAFTCKRKYVOREGRA 410
 DB 451 LITRRGGEDDKTEIFRRPGGDMRNWSELYKYRVNKLIEPLGVAFTCKRKYVOREGRA 510
 QY 411 VGIGAVFIFGFGAGSTWGAASMTLTVQARLLLSGIYVOOONNLJLAIEAQQRMLOLTYWG 470
 DB 511 VGIGAVFIFGFGAGSTWGAASMTLTVQARLLLSGIYVOOONNLJLAIEAQQRMLOLTYWG 570
 QY 471 IKQIOLARVLAVERVLDGDOOLGIMWCSGKLICTTAVPNNASNKSJDRINNNMTWME 530
 DB 571 IKQIOLARVLAVERVLDGDOOLGIMWCSGKLICTTAVPNNASNKSJDRINNNMTWME 630
 QY 531 REIDNYSBITTLIESONQOEKNEOELLEDDKMASLWNPEDITWLMY 579
 DB 631 REIDNYSBITTLIESONQOEKNEOELLEDDKMASLWNPEDITWLMY 679

RESULT 5
 ID AAW88113
 XX AAW88113 standard; Protein: 855 AA.
 AC AAW88113;
 XX
 DT 09-APR-1999 (first entry)
 XX
 DE Env protein of the BA-L strain of Human immunodeficiency virus type 1.
 XX
 KW HIV-1; HIV-1 strain BA-L; env protein; vaccine;
 XX immunotherapy; HIV infection; immunogen; HIV-1 diagnosis.
 OS Human immunodeficiency virus type 1.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 11
 FT
 XX US5869313-A.
 PN 09-FEB-1999.
 XX 14-MAY-1996; 96US-0647714.
 PF 17-OCT-1990; 90US-0599491.
 XX 25-FEB-1993; 93US-0022835.
 PR 15-FEB-1995; 95US-0388809.
 PR 14-MAY-1996; 96US-0647714.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA
 XX Franchini G, Gallo RC, Gartner S, Lori FC, Markham PD;
 PI Popovic M, Reitz MS;
 XX WPI: 1999-152779/13.
 DR N-PSDB; AAX04767.
 DR
 XX DNA encoding env protein of the human immune deficiency virus
 PT isolate BA-L - useful for producing protein for use in vaccines, as
 PT assay reagent and to generate antibodies
 XX
 PS Example 1; Fig 9A-C; 87pp; English.
 XX
 CC The present sequence represents the envelope protein of the BA-L
 CC (ATCC 40890) strain of Human immunodeficiency virus type 1 (HIV-1)
 CC strain MN-ST1. BA-L is more typical of United States isolates of
 CC HIV-1 than previously known strains. Recombinant, complete env protein
 CC of the BA-L strain is used as a vaccine component and for immunotherapy
 CC of existing HIV infections, to detect HIV-specific antibodies, e.g. in
 CC donated blood, and as an immunogen to raise specific antibodies, for
 CC HIV-1 diagnosis.
 CC
 XX
 SQ Sequence 855 AA;
 Query Match 89.5%; Score 2800.5; DB 20; Length 855;
 Best Local Similarity 81.2%; Pred. No. 1e-129;
 Matches 527; Conservative 20; Mismatches 31; Indels 71; Gaps 2;
 QY 2 EKLWVTVYGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTE 61
 DB 31 EKLWVTVYGVVWKEATTTLFCASDRKAYDTEVHNWATHACVPTDNPQEVLENTE 90
 QY 62 HFNWKNWVQWQEDIIISLDQSLKPCVKLTPLC----- 96
 DB 91 NFNWKNWVQWQEDIIISLDQSLKPCVKLTPLC----- 150
 QY 97 --GAG----- 110
 DB 151 MVGGGEMKNCSEFNTTNRGKVKQEVAFYKLDIAPIDNNSNNRRLISCTSVITQACP 210
 QY 111 KISFEPIPIHYCAPAGFAILKNDKTEGKPCKNVSTVQCTHGIRPVVSTQLLNGSLA 170

```

Db      211 KSFSEPIPIHYCAPAGFALICKCKKFKNGKPCPTNVSTVQCTHGIRPVSTOLLNGSLA 270
Qy      171 EEEVIRSDNFTNNKTIIVOLKESVEINCTRPNNTSRSHIHGPRARVYTGELIGDR 230
Db      271 EEEVIRSANFADNAKVIIVOLNESVEINCTPNNTSRSHIHGPRARVYTGELIGDR 330
Qy      231 QAHCHISRAKAMDLLKQIVIKLREOPENKTIIVFNHSSGGDEIYVMSFNGCGEPFYCNST 290
Db      331 QAHCHLSRAKAMDLLINKIYIKLREOPENKTIIVFKHSSGGDEIYVMSFNGCGEPFYCNST 390
Qy      291 QLFNSTWNNNTGSSNNTBENTITLPCRKOIINMGEVAKYAPPIRQIRCSNITGL 350
Db      391 QLFNSTWNNNTGSSNNTBENTITLPCRKOIINMGEVAKYAPPIRQIRCSNITGL 450
Qy      351 ILTRGGIINENGETEIPFGGDMRDWMSSELYKKYKVKLEPIGVAPTKRRRVQEKKA 410
Db      451 ILTRGGIINENGETEIPFGGDMRDWMSSELYKKYKVKLEPIGVAPTKRRRVQEKKA 510
Qy      411 VGIGAVFLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOONNLRAIEAQOHLQLTWVG 470
Db      511 VGIGAVFLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOONNLRAIEAQOHLQLTWVG 570
Qy      471 IKQLOARVLAVERYLADQQLIGWCGSKLICCTAVPANAASNSLDRIMNMTWMEWE 530
Db      571 IKQLOARVLAVERYLADQQLIGWCGSKLICCTAVPANAASNSLDRIMNMTWMEWE 630
Qy      531 REIDNTSIEITLIESONOQEKNEBELLEDKASLSLMMWPDITNWLWY 579
Db      631 REINNTSIIYSLIESONOQEKNEBELLEDKASLSLMMWPDITNWLWY 679

RESULT 6
AAB82761
ID      AAB82761 standard, Protein, 883 AA.
XX      AAB82761;
XX      29-OCT-2001. (first entry)
XX      DE      Ancestral HIV-1 group M, subtype B gp160 protein.
XX      KW      HIV-1; env gene; gp160 protein; ancestral gene; phylogeny; vaccine.
XX      OS      Human immunodeficiency virus type 1.
XX      PN      W0200160838-A2.
XX      23-AUG-2001.
PF      16-FEB-2001; 2001WO-US05288.
XX      PR      18-FEB-2000; 2000US-0183659.
XX      (UNITM ) UNIV WASHINGTON.
XX      PA      Mullins JI, Rodrigo AG, Learn GH, Li F;
XX      PI      WPI: 2001-536565/59.
XX      DR      N-PSDB; AAH26468.
PT      Preparing an ancestral viral amino acid sequence useful as a vaccine
PT      comprises determining a recent ancestor of a circulating virus by
PT      maximum likelihood phylogeny analysis -
XX      Claim 8; Page 54; 89pp; English.
XX      The present sequence is that of an ancestral HIV-1 group M, subtype
XX      B gp160 (env gene product) sequence. The invention provides
XX      compositions and methods for determining ancestral viral gene
XX      sequences and ancestral viral protein sequences for highly diverse
XX      viruses, such as HIV-1. The methods use samples of circulating
XX      viruses to determine an ancestral viral sequence by maximum

```

```

CC      likelihood phylogeny analysis. In the present case, the ancestral
CC      HIV-1 subtype B env sequence (see AAH26468) was determined using 38
CC      subtype B sequences (obtained from 9 different countries) and 3
CC      subtype D (outgroup) sequences. The distances between this ancestral
CC      viral sequence and circulating strains used to determine it were on
CC      average 12.3% (range: 8.0-21.0%) while the available specimens were
CC      17.3% different from each other (range: 13.3-23.2%). Thus, the
CC      ancestor sequence was, on average, more closely related to any given
CC      circulating virus than to any other variant. The ancestral gp160
CC      sequence included a wide variety of immunogenically active peptides
CC      when processed for antigen presentation; nearly all known subtype B
CC      CTL epitope consensus amino acids were represented. Thus, an
CC      immunogenic composition to this subtype B ancestor protein will
CC      elicit broad neutralising antibody against HIV-1 isolates of the
CC      same subtype, and will also elicit a broad cellular response
CC      mediated by antigen-specific T-cells. A claimed vaccine composition
CC      comprises a viral ancestor protein or its immunogenic fragment,
CC      especially one derived from the HIV-1 group M subtype B gp160
CC      ancestral protein.
XX      SQ      Sequence      883 AA;
Qy      Query Match      88.9%; Score 2782; DB 22; Length 883;
Db      Best Local Similarity 79.2%; Pred. No. 8.4e-129;
Qy      Matches 537; Conservative 14; Mismatches 27; Indels 100; Gaps 7;
Qy      2 EKLMTVYVGVPMVEAATTLPCASDAKAYDEVNMAWATHCVPTDPPOVLENTYE 61
Db      31 EKLMTVYVGVPMVEAATTLPCASDAKAYDEVNMAWATHCVPTDPPOVLENTYE 90
Qy      62 HFNMAKNNMVEQMOEDISLMDQSLKPCVKLTPLC----- 96
Db      91 HFNMAKNNMVEQMHEDIISLMDQSLKPCVKLTPLCVTLNCTDLDATNATNTNSSATTNT 150
Qy      97 ---GAG----- 99
Db      151 TSSGGTMEGEKGEIKNCSFVNTTISRDMQKEYALFYKLDVVPIDNDNNNTNTSYRL 210
Qy      100 --CDTSVITQACPKSPFPIPIHYCAPAGFALICKNDKTFNKGKQXNVSTVQCHGRP 157
Db      211 INCNTSVITQACPKSPFPIPIHYCTPAGFALICKNDKTFNKGKQXNVSTVQCHGRP 270
Qy      158 VVSTOLLNGSLAESEVIRSDNFTNNKTIIVOLKESVEINCTRPNNTSRSHIHGPR 217
Db      271 VVSTOLLNGSLAESEVIRSDNFTNNKTIIVOLKESVEINCTRPNNTSRSHIHGPR 330
Qy      218 AFYTTGELIGDIRQAHCHISRAKAMDLLKQIVIKLREOPENK--TIVFNHSSGGDEIYV 275
Db      331 ALYATGKILIGDIRQAHCHISRAKAMDLLKQIVIKLREOPENKTIIVFNHSSGGDEIYV 390
Qy      276 HSFNCGGEFFYCNSTQLTNNTW-----NNTBESNNT--EGNTITLPCRKOIINMW 326
Db      391 HSFNCGGEFFYCNSTQLTNNTWTFNFGTWGNNNTESNNAADNDITLPCRKOIINMW 450
Qy      327 EYKAMVAPPIRQIRCSNITGLLTRDGGINEN----GTEIPFGGDMRDNRSELY 382
Db      451 EYKAMVAPPIRQIRCSNITGLLTRDGGINENNTTDIRIPFGGDMRDNRSELY 510
Qy      383 KYKVKIEPLGVAPTKRRRVQEKRAVG--IGAVFLGFLGAAGSTWGAASMTLTVQARL 441
Db      511 KYKVKIEPLGVAPTKRRRVQEKRAVGILGAVFLGFLGAAGSTWGAASMTLTVQARL 570
Qy      442 LLSGIVQOONNLRAIEAQOHLQLTWVGIKQLOARVLAVERYLADQQLIGWCGSKLI 501
Db      571 LLSGIVQOONNLRAIEAQOHLQLTWVGIKQLOARVLAVERYLADQQLIGWCGSKLI 630
Qy      CTAVPANASNSKSLDRIMNMTWMEWEREIDNTSIEITLIESONOQEKNEBELLE 561
Db      631 CTAVPANASNSKSLDRIMNMTWMEWEREIDNTSIEITLIESONOQEKNEBELLE 690
Qy      562 DKWASLMMWPDITNWLWY 579
Db      691 DKWASLMMWPDITNWLWY 708

```

RESULT 7

AA67724 standard; Protein: 850 AA.

AA67724;

25-MAR-2003 (updated)
07-SEP-1995 (first entry)

gp120 from the HIV GNE8 isolate.

HIV; human immunodeficiency virus; gp120; glycoprotein;
GNE clone; GNE8 isolate.

Human immunodeficiency virus type 1.

MO9428929-A1.

22-DEC-1994.

07-JUN-1994; 94WO-US06036.

07-JUN-1993; 93US-0072833.

(GETH) GENENTECH INC.

Berman PW, Nakamura GR;

WPI; 1995-036112/05.

N-PSDB; AA076018.

Use of HIV gp 120 polypeptide(s) - for developing probes for the
analysis, prevention and therapy of HIV infection

Claim 23; Page 27-30; 108pp; English.

This protein is the gp120 (envelope glycoprotein) from the GNE8 isolate
of HIV-1 (human immunodeficiency virus type 1). There are neutralising
epitopes in the V2 and C4 domains of gp120, in addition to the
neutralising epitopes in the V3 domain. Although the amino acid sequences
of the neutralising epitopes in the V2, V4 and C4 domains are variable,
the amount of variation is highly constrained. This facilitates the
design of HIV subunit vaccines that can induce antibodies that neutralise
the most common HIV strains for a given geographic region. This invention
provides a multivalent gp120 subunit vaccine where the gp120 present in
the vaccine is from at least two HIV isolates which have different amino
acid sequences for a neutralising epitope in these regions.
(Updated on 25-MAR-2003 to correct PN field.)

Sequence . 850 AA;

Query Match 88.7%; Score 2776.5; DB 16; Length 850;
Best Local Similarity 81.5%; Pred. No. 1.5e-128;
Matches 528; Conservative 19; Mismatches 28; Indels 73; Gaps 5;

QY 2 EKLMTVYGVVWVKEATTTTCASDAKAYDTEVHNWATHACVTDNDPOEVLENTE 61
DB 31 EKLMTVYGVVWVKEATTTTCASDAKAYDTEVHNWATHACVTDNDPOEVLENTE 90
QY 62 HFNWKNMNVEMOEDISLMDQSLKPCVKLTPLCGA----- 98
DB 91 NFNWKNMNVEMOEDISLMDQSLKPCVKLTPLCGA----- 150
QY 99 -----GCDTSVTQAQCPKISFEP 116
DB 151 EIKNCSFVNTSIRDKMKNVAFYKLDVPIIDNDNTSYRLISCTSVITQAQCPKISFEP 210
QY 117 IPIHYCAAGPAFLKCNKTPGKPCPKVSVQCTHGRIPVSVQLLNGSLAEVVI 176
DB 211 IPIHYCAAGPAFLKCNKTPGKPCPKVSVQCTHGRIPVSVQLLNGSLAEVVI 270

QY 177 RSDNFTNNAKTIIVOLKESVEINCTRPNNNTKRSIHIGPRAFYTTGIIIGDIRQAHNT 236
DB 271 RSNFSDNNAKTIIVOLKESVEINCTRPNNNTKRSIHIGPRAFYATGIIIGDIRQAHNT 330
QY 237 SRKMNUTLKQIVYKLEQEPENKTIIVNHSAGDPPEIWMHSFNGGFFFCNSQLNST 296
DB 331 SSTKMNTLKQIVYKLEHFP-NKTIIVNHSAGDPPEIWMHSFNGGFFFCNTPLNST 389
QY 297 WN-----NNTGSSNNTGNTITLPCRIKOIINMWOEGKAMVAPPIRGQIRCSNITGL 351
DB 390 WNYTYTNNTGSSNDT-GRNITLQCRIKOIIINMWOEGKAMVAPPIRGQIRCSNITGL 448
QY 352 LTRDGINENGTEIFRPGGDMRDMNRSSELYKKVVIIEPLGVAPTYCKRRVYQREKRAV 411
DB 449 LTRDGG-NNSETTEIFRPGGDMRDMNRSSELYKKVVIIEPLGVAPTYCKRRVYQREKRAV 507
QY 412 GIGAVPLGFLGAAGSTMGAAASMTLTVOARLLSGIVQOQNNLRAIEAQPMLOLTWGI 471
DB 508 GIGAVPLGFLGAAGSTMGAAASMTLTVOARLLSGIVQOQNNLRAIEAQPMLOLTWGI 567
QY 472 KOLQARVLAVERYLGDQOLGIMWCSGKLICTTAVPNNASMSKSLDRINNNMTMEMER 531
DB 568 KOLQARVLAVERYLKDQOLGIMWCSGKLICTTAVPNNASMSKSLDKINDNTMEMER 627
QY 532 EIDNVTSEIYTLIESQNOQEKNEOELLEDKWASLWNPEDITWLMY 579
DB 628 EIDNVTSLIYSLIESQNOQEKNEOELLEDKWASLWNPEDITWLMY 675

RESULT 8

AAU75156 standard; Protein: 619 AA.

AAU75156;

08-MAY-2002 (first entry)

N-terminal mutant of modified HIV-1 SF162DV2 gp140 envelope protein.

Human immunodeficiency virus type 1; HIV-1; gp140 envelope protein;
second hypervariable region; V2; SF162deltaV2; SF162DV2; AIDS;
acquired immunodeficiency syndrome; anti-HIV; mutant; muten; immunogen.
Human immunodeficiency virus type 1 strain SF162 (clade B).
Synthetic.

Key Location/Qualifiers
FT Misc-difference 512
FT note= "Encoded by CAGCTC"

MO200200250-A2.

03-JAN-2002.

27-JUN-2001; 2001WO-US20483.

27-JUN-2000; 2000US-214608P.

26-JUN-2001; 2001US-0891609.

(AARO-) AARON DIAMOND AIDS RES CENT.

Stamatatos L, Barnett S, Shrivastava I;

WPI; 2002-130836/17.

N-PSDB; AAS15499.

Immunizing an animal and eliciting an immune response against
heterologous HIV-1 in an animal, involves administering an immunogen
comprising modified HIV-1 envelope protein, or DNA or virus encoding
the protein -
Claim 5; Fig 17; 62pp; English.

CC The present invention relates to methods for immunising an animal,
 CC preferably human, against heterologous human immunodeficiency virus
 CC type 1 (HIV-1), and eliciting a heterologous immune response to HIV-1
 CC in that animal. The method comprises administering an immunogen having
 CC at least one modified HIV-1 envelope protein with a deletion in the
 CC second hypervariable (V2) region or its fragment, or DNA or virus
 CC encoding the modified HIV-1 envelope protein. The invention discloses
 CC the polynucleotide and amino acid sequences for the full-length HIV-1
 CC SF162deltaV2 (SF162DV2) gp140 envelope protein, and, for an N-terminal
 CC mutant of SF162DV2. The immunogens of the invention are useful for
 CC immunising an animal against heterologous HIV-1 strains by eliciting
 CC neutralising antibodies or protective antibodies in the animal. The
 CC method can be used to treat acquired immunodeficiency syndrome (AIDS).
 CC The present sequence represents the N-terminal mutant (deletion of
 CC amino acids 1-27) of modified HIV-1 SF162DV2 gp140 envelope protein.

SO Sequence 619 AA;

Query Match 88.6%; Score 2771; DB 23; Length 619;

Best Local Similarity 85.6%; Pred. No. 2, 2e-128;

Matches 529; Conservative 16; Mismatches 31; Indels 42; Gaps 6;

QY 1 VEKLVVTVYVGVVWKEATTLTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENT 60
 DB 3 VEKLVVTVYVGVVWKEATTLTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENT 62
 QY 61 EHFNNMKNMVEQWQEDIIISLMDQSLKPCVKLTPLC----- 96
 DB 63 EHFNNMKNMVEQWQEDIIISLMDQSLKPCVKLTPLC----- 122
 QY 97 -----GAG-----CDTSVITQACPKISFEPPIHYCAPAGPAIILKNDKTFNGKGP 142
 DB 123 GEIKNCSFKVAGKGLNCTSVITQACPKVSEFPIHYCAPAGPAIILKNDKTFNGKGP 182
 QY 143 CKNVSTVQCTHGIRPVVSTQLLNGSLAEVEVIRSDNFTNNAKTIIVQLKSEVINCTR 202
 DB 183 CTNVSTVQCTHGIRPVVSTQLLNGSLAEVEVIRSDNFTNNAKTIIVQLKSEVINCTR 242
 QY 203 PNNNTRKSIHIGPRAFYTTGELIGIRQAHGNISAKKNDTLKOIVILRQFENKTI 262
 DB 243 PNNNTRKSTTIGPRAFYATGDIIGIRQAHGNISAKKNDTLKOIVILRQFENKTI 302
 QY 263 FNNHSGGDEIYVHSGFCGSEFFYCNSTQLFNSVNNNTGSGNTEGNTITLPCRIKOI 322
 DB 303 FNNHSGGDEIYVHSGFCGSEFFYCNSTQLFNSVNNNTGSGNTEGNTITLPCRIKOI 360
 QY 323 NNMQVGVKAMVAPPIRGQIRCSNITGLLLTDDGIN-ENGTEIFRPGGDMRDWRSEL 381
 DB 361 NNMQVGVKAMVAPPIRGQIRCSNITGLLLTDDGKEISNTTEIFRPGGDMRDWRSEL 420
 QY 382 YKXKVKIETPLGVAPTKRRVQREKRAVVGAVELGFLGAGSGTGAASMTLVQAAL 441
 DB 421 YKXKVKIETPLGVAPTKRRVQREKRAVVGAVELGFLGAGSGTGAASMTLVQAAL 480
 QY 442 LLSGIVQOQNNLLRAIEAQRMQLQTVWGIKQLQARVLAVERVLDQQLIGWCSGKLI 501
 DB 481 LLSGIVQOQNNLLRAIEAQHLLQLTWGIK-LQARVLAVERVLDQQLIGWCSGKLI 539
 QY 502 CCTAVPMAWMSNKSILDRIMNMTWMEWREIDNTYSIYLLIEESQOEKNEDELLEL 561
 DB 540 CCTAVPMAWMSNKSILDRIMNMTWMEWREIDNTYLLIYLLIEESQOEKNEDELLEL 599
 QY 562 DKWASLAWMFDITNLMWY 579
 DB 600 DKWASLAWMFDITNLMWY 617

RESULT 9
 AAU75155 ID AAU75155 standard; Protein; 646 AA.
 XX AC AAU75155;
 XX

DT 08-MAY-2002 (first entry)
 XX Modified full-length HIV-1 SF162DV2 gp140 envelope protein.
 DE
 XX Human immunodeficiency virus type 1; HIV-1; gp140 envelope protein;
 KW second hypervariable region; V2; SF162deltaV2; SF162DV2; AIDS;
 KW acquired immunodeficiency syndrome; anti-HIV; mutant; mucin; immunogen.
 XX
 OS Human immunodeficiency virus type 1 strain SF162 (clade B).
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 539
 FT /note= "Encoded by CAGCTC"
 XX
 PN WO200200250-A2.
 XX
 PD 03-JAN-2002.
 XX
 PD 27-JUN-2001; 2001WO-US20483.
 PF
 XX 27-JUN-2000; 2000US-214608P.
 PR 26-JUN-2001; 2001US-0891609.
 XX
 PA (AARO-) AARON DIAMOND AIDS RES CENT.
 XX
 PI Stamatatos L, Barnett S, Shrivastava I;
 XX
 DR WPI; 2002-130836/17.
 DR N-PSDB; AAS15498.
 XX
 PT Immunizing an animal and eliciting an immune response against
 PT heterologous HIV-1 in an animal, involves administering an immunogen
 PT comprising modified HIV-1 envelope protein, or DNA or virus encoding
 PT the protein -
 XX
 PS Claim 5; Fig 16; 62pp; English.
 XX
 CC The present invention relates to methods for immunising an animal,
 CC preferably human, against heterologous human immunodeficiency virus
 CC type 1 (HIV-1), and eliciting a heterologous immune response to HIV-1
 CC in that animal. The method comprises administering an immunogen having
 CC at least one modified HIV-1 envelope protein with a deletion in the
 CC second hypervariable (V2) region or its fragment, or DNA or virus
 CC encoding the modified HIV-1 envelope protein. The invention discloses
 CC the polynucleotide and amino acid sequences for the full-length HIV-1
 CC SF162deltaV2 (SF162DV2) gp140 envelope protein, and, for an N-terminal
 CC mutant of SF162DV2. The immunogens of the invention are useful for
 CC immunising an animal against heterologous HIV-1 strains by eliciting
 CC neutralising antibodies or protective antibodies in the animal. The
 CC method can be used to treat acquired immunodeficiency syndrome (AIDS).
 CC The present sequence represents the full-length modified HIV-1
 CC SF162DV2 gp140 envelope protein.
 CC
 XX
 SO Sequence 646 AA;

Query Match 88.6%; Score 2771; DB 23; Length 646;
 Best Local Similarity 85.6%; Pred. No. 2, 3e-128;

Matches 529; Conservative 16; Mismatches 31; Indels 42; Gaps 6;

QY 1 VEKLVVTVYVGVVWKEATTLTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENT 60
 DB 30 VEKLVVTVYVGVVWKEATTLTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENT 89
 QY 61 EHFNNMKNMVEQWQEDIIISLMDQSLKPCVKLTPLC----- 96
 DB 90 EHFNNMKNMVEQWQEDIIISLMDQSLKPCVKLTPLC----- 149
 QY 97 -----GAG-----CDTSVITQACPKISFEPPIHYCAPAGPAIILKNDKTFNGKGP 142
 DB 150 GEIKNCSFKVAGKGLNCTSVITQACPKVSEFPIHYCAPAGPAIILKNDKTFNGKGP 209
 QY 143 CKNVSTVQCTHGIRPVVSTQLLNGSLAEVEVIRSDNFTNNAKTIIVQLKSEVINCTR 202


```

Db 210 CTNVTAVCTGIRRVSTOLLNGLSLAEGVIRSENFNTNNAKTIIVQLKESVEINCTR 269
Qy 203 PNNNRKSIHIGPGAFYTTGELIGDIOAHCHNISRAKWNLTQIYVLEKQENKTIIV 262
Db 270 PNNNRKSIHIGPGAFYTTGELIGDIOAHCHNISRAKWNLTQIYVLEKQENKTIIV 329
Qy 263 FNNHSGGDEPEIVMHSFNGCGEFPYCNSTOLFNTNNNTGSSNTEGNTITLPCRKOI 322
Db 330 FKQSSGDEPEIVMHSFNGCGEFPYCNSTOLFNTNNNTGSSNTEGNTITLPCRKOI 387
Qy 323 NMQGVKAMAPPIRGQIRGSSNTITGLTRDGIN-ENGTEIFRPGGDMRDMRSEL 381
Db 388 NRMQGVKAMAPPIRGQIRGSSNTITGLTRDGIN-ENGTEIFRPGGDMRDMRSEL 447
Qy 382 YKYVVKIEPIGVAPTKRRRVQEKRAVIGAFELGPAAGSTMAAAMTLTVQARL 441
Db 448 YKYVVKIEPIGVAPTKRRRVQEKRAVIGAFELGPAAGSTMAAAMTLTVQARL 507
Qy 442 LLSGIVQOQNNLRAIEAQRMLOLTWGIKOLQARVLAVERYLQDQOLLGIWCGSGKLI 501
Db 508 LLSGIVQOQNNLRAIEAQRMLOLTWGIKOLQARVLAVERYLQDQOLLGIWCGSGKLI 566
Qy 502 CCTAVPMAASWSNKSLEDRIMNNMTWMEWEREIDNTSEIYTLIESQNOEKNEQELLE 561
Db 567 CTTAVPMAASWSNKSLEDRIMNNMTWMEWEREIDNTSEIYTLIESQNOEKNEQELLE 626
Qy 562 DKMASLNNWPDITWMLWY 579
Db 627 DKMASLNNWPDITWMLWY 644

```

RESULT 10

```

AA97073
ID AA97073 standard; protein; 847 AA.
XX
AC AA97073;
XX
DT 31-OCT-2000 (first entry)
XX
DE Variant HIV-1 SFI62 Env gp160.
XX
KW HIV-1; vaccine; anti-HIV; immunogenic; modified; envelope; Env;
KM CD4 binding region; V1/V2 loop; bridging sheet.
XX
OS Human immunodeficiency virus type 1 isolate SFI62.
XX
MO200039303-A2.
XX
06-JUL-2000.
XX
PE 30-DEC-1999; 99WO-US31272.
XX
PR 31-DEC-1998; 98US-0114495.
PR 29-SEP-1999; 99US-0156670.
XX
PA (CHIR ) CHIRON CORP.
XX
PI Barnett S, Hartog K, Martin E;
XX
WP1; 2000-465745/40.
XX
PT Novel modified HIV Env polypeptides useful as immunizing agents and for
PT preparing a vaccine to elicit an immune response against a broad range
PT of HIV subtypes
XX
PS Claim 5; Page 115-117; 139pp; English.
XX
CC Novel immunogenic modified human immunodeficiency virus (HIV) envelope
CC (Env) polypeptides having an amino acid deleted or replaced in the region
CC corresponding to residues 420-436 or 119-123 and 199-210 relative to
CC isolate HXB-2 are disclosed. The modified Env polypeptide is based on HIV
CC strain SFI62, with numbering relative to isolate HXB-2. The Env

```

CC polypeptides are modified so as to expose at least part of the CD4
 CC binding region. The modified HIV Env polypeptides, coding polynucleotides
 CC and constructs, further comprising an adjuvant, are used for inducing an
 CC immune response in an individual. The method involves administering a
 CC first composition comprising a polynucleotide encoding the Env
 CC polypeptide in a priming step and administering a second composition
 CC comprising a modified Env polypeptide as a booster in an amount
 CC sufficient to induce an immune response in the individual. The first
 CC and/or second composition further comprises an adjuvant (claimed). The
 CC intracellularly produced Env polypeptides can be used for a number of
 CC diagnostic and therapeutic purposes to determine the presence of reactive
 CC antibodies/and or Env proteins in a biological sample to aid in the
 CC diagnosis of HIV infection or disease status or as measure of response to
 CC immunization.

XX Sequence 847 AA;

Query Match 88.3%; Score 2764; DB 21; Length 847;
 Best Local Similarity 81.7%; Pred. No. 6.2e-128;
 Matches 527; Conservative 16; Mismatches 34; Indels 68; Gaps 4;

```

Qy 1 VEKLVTVYVGVVPWKATTTLPASDAKAYDREHNVMAATHACVPDPDPOEVLENTV 60
Db 30 VEKLVTVYVGVVPWKATTTLPASDAKAYDREHNVMAATHACVPDPDPOEVLENTV 89
Qy 61 EHFNNMKNWVEQKQEDIIISLMDQSLKPCVLTPLCGA----- 98
Db 90 EHFNNMKNWVEQKQEDIIISLMDQSLKPCVLTPLCTYHCTNKNATNTKSNWKMEDR 149
Qy 99 -----GDTSVITQACPKISFE 115
Db 150 GEIKNSCFKVTTSIRNKKQKAYALFYKLDVVPIDNDNTSYKLINCNTSVITQACPKISFE 209
Qy 116 PIPHYCAPAGFALKNDKTFNGKPGCKNASTVQCTHGRIPVYSTOLLNGLSLAESEVY 175
Db 210 PIPHYCAPAGFALKNDKTFNGKPGCKNASTVQCTHGRIPVYSTOLLNGLSLAESEVY 269
Qy 176 IRSNFTNNAKTIIVQLKESVEINCTPRNNTRKSIHIGPRAFYTTGELIGDIOAHCHN 235
Db 270 IRSNFTNNAKTIIVQLKESVEINCTPRNNTRKSIHIGPRAFYTTGELIGDIOAHCHN 329
Qy 236 ISRAKNDTLKOIYVLEKQENKTIIVFNNHSGGDEPEIVMHSFNGCGEFPYCNSTOLFNS 295
Db 330 ISRAKNDTLKOIYVLEKQENKTIIVFNNHSGGDEPEIVMHSFNGCGEFPYCNSTOLFNS 389
Qy 296 TWNNNTGSSNTEGNTITLPCRKOIINMMEVGRKAMAPPIRGQIRGSSNTITGLTRD 355
Db 390 TWNNNTGSSNTEGNTITLPCRKOIINMMEVGRKAMAPPIRGQIRGSSNTITGLTRD 447
Qy 356 GGIN-ENGTEIFRPGGDMRDMRSELYKYVVKIEPIGVAPTKRRRVQEKRAVIG 414
Db 448 GGIN-ENGTEIFRPGGDMRDMRSELYKYVVKIEPIGVAPTKRRRVQEKRAVIG 507
Qy 415 AVFLGFLGAGSTNGAASMTLTVQARLLSGIVQOQNNLRAIEAQRMLOLTWGIKOL 474
Db 508 AVFLGFLGAGSTNGAASMTLTVQARLLSGIVQOQNNLRAIEAQRMLOLTWGIKOL 567
Qy 475 QARVLAVERYLQDQOLLGIWCGSGKLICTTAVPMAASWSNKSLEDRIMNNMTWMEWEREID 534
Db 568 QARVLAVERYLQDQOLLGIWCGSGKLICTTAVPMAASWSNKSLEDRIMNNMTWMEWEREID 627
Qy 535 NYTSEIYTLIESQNOEKNEQELLEDKMASLNNWPDITWMLWY 579
Db 628 NYTSEIYTLIESQNOEKNEQELLEDKMASLNNWPDITWMLWY 672

```

RESULT 11

```

ABU6565
ID ABU6565 standard; Protein; 842 AA.
XX
AC ABU6565;
XX
DT 22-MAY-2003 (first entry)

```


CC useful in gene therapy and DNA immunisation applications. ABL39942 to
CC ABL40054 and ABB06204 to ABB06215 represent sequences used in the
CC exemplification of the present invention.

XX Sequence 842 AA;

Query Match 87.9%; Score 2749; DB 23; Length 842;
Best Local Similarity 81.2%; Pred. No. 3.3e-127;
Matches 524; Conservative 16; Mismatches 37; Indels 68; Gaps 4;

```

QY 1 VEKLVTVYGVVWPKKATTTLLFCASDAKAYDEVHANNWATHACVPTDPNPOEVLENTV 60
DB 25 VEKLVTVYGVVWPKKATTTLLFCASDAKAYDEVHANNWATHACVPTDPNPOEVLENTV 84
QY 61 EHFNNKNNWEOQOEDIIISLMDOSLKPCKVLTPLCGA----- 98
DB 85 EHFNNKNNWEOQOEDIIISLMDOSLKPCKVLTPLCGA----- 144
QY 99 -----GCDTSVITQACPCKISFE 115
DB 145 GEIKNCSFKVTTIRNKKQKEYALFYKLDVVPIDNDNTSYKLINCNTSVTTQACPKVSFE 204
QY 116 PIPHYCAPGAFALCKNDKTPNGKPCNKANSTVOCTHGIRPVYSTQLLNGSLAEEV 175
DB 205 PIPHYCAPGAFALCKNDKTPNGKPCNKANSTVOCTHGIRPVYSTQLLNGSLAEEV 264
QY 176 IRSNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGPRAFYTTGELIGDIROACHN 235
DB 265 IRSNFTNNAKTIIVOLKESVEINCTRPNNNTRKSTTIGPRAFYATGDIIGDIROACHN 324
QY 236 ISRAKNDTLKQIYIKLREOFENKTIIVNHSFGDPEIVMHSFNGGGEFYCNSTOLFNS 295
DB 325 ISGEKNNNTLKQIYIKLQAFGKNTIVFKQSSGGDEIVMHSFNGGGEFYCNSTOLFNS 384
QY 296 TWANNTEGSNNTEGNTITLPCRKOIINMOEYKAMVAPPIRGDIRCSSNITGLLTRD 355
DB 385 TW-NNITGPNNTNG-TITLPCRKOIINMOEYKAMVAPPIRGDIRCSSNITGLLTRD 442
QY 356 GGIN-ENGTEIFRPGGGDMRDMRSELKYKVKVKEIPLGVAPTCKRRVVOREKRAVGIG 414
DB 443 GGEKISNTTEIFRPGGGDMRDMRSELKYKVKVKEIPLGVAPTCKRRVVOREKRAVGIG 502
QY 415 AVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOQNNLLRAIEAOQRLQLTWGIKQL 474
DB 503 AMFLGFLGAAGSTMGASMTLTVQARQLLSGIVQOQNNLLRAIEAOQRLQLTWGIKQL 562
QY 475 QARVLAVERYLGDQQLLGITWGSGLICTTAVPWNASWSKSLDRINNNMTWMEWEREID 534
DB 563 QARVLAVERYLKDQQLLGITWGSGLICTTAVPWNASWSKSLDRINNNMTWMEWEREID 622
QY 535 NYTSEIYTLIESONQOEKNEOELELDPMASLMMWPDITNLMWY 579
DB 623 NYTULYTLIESONQOEKNEOELELDPMASLMMWPDISKLMWY 667

RESULT 13
AAP80967
ID AAP80967 standard; protein; 851 AA.
AC AAP80967;
XX
XX 25-MAR-2003 (updated)
DT 15-NOV-1990 (first entry)
XX
DB HIV protein HT6.
XX HIV; HT6; gp160; envelope protein; RF; AIDS.
XX Human immunodeficiency virus variant RF.
XX EP272858-A.
XX 29-JUN-1988.
PD

```

```

XX 14-DEC-1987; 87EP-0310967.
XX PF 31-AUG-1987; 87US-0091481.
XX PR 15-DEC-1986; 86US-0941111.
XX (REPK) REPLIGEN CORP.
XX
XX Rusche J, Lynn D, Carson H, Putney S, Jellis CL;
XX WPI; 1988-176944/26.
XX DR N-PSDB; AAN80948.
XX
XX The sequence is the result of cloning a hybrid envelope gene from
XX HIV variants BH10 and RF. A central portion of the RF gene was
XX used, the rest being from the distantly related variant BH10. The
XX resulting clone, pACH6 produces a hybrid gp 160 envelope protein
XX with novel immunological and antigenic characteristics. It may
XX be used to as a vaccine and for diagnosis and therapy of AIDS.
XX See also AAP80966.
XX (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 851 AA;
XX
XX Query Match 87.4%; Score 2735.5; DB 9; Length 851;
XX Best Local Similarity 79.9%; Pred. No. 1.5e-126;
XX Matches 516; Conservative 33; Mismatches 29; Indels 69; Gaps 4;

QY 2 EKLWTVYGVVWPKKATTTLLFCASDAKAYDEVHANNWATHACVPTDPNPOEVLENTV 61
DB 32 EKLWTVYGVVWPKKATTTLLFCASDAKAYDEVHANNWATHACVPTDPNPOEVLENTV 91
QY 62 HFNMMKNNWEOQOEDIIISLMDOSLKPCKVLTPLCGA----- 98
DB 92 HFNMMKNNWEOQOEDIIISLMDOSLKPCKVLTPLCGA----- 151
QY 99 -----GCDTSVITQACPCKISFE 115
DB 152 GEIKNCSFNNSTIRGKQKEYAFYKLDIIPIDNDTSTYKLSNCTSVITQACPCKVSFE 211
QY 116 PIPHYCAPGAFALCKNDKTPNGKPCNKANSTVOCTHGIRPVYSTQLLNGSLAEEV 175
DB 212 PIPHYCAPGAFALCKNDKTPNGKPCNKANSTVOCTHGIRPVYSTQLLNGSLAEEV 271
QY 176 IRSNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGPRAFYTTGELIGDIROACHN 235
DB 272 IRSNFTNNAKTIIVOLKESVEINCTRPNNNTRKSTTIGPRAFYATGDIIGDIROACHN 331
QY 236 ISRAKNDTLKQIYIKLREOFENKTIIVNHSFGDPEIVMHSFNGGGEFYCNSTOLFNS 295
DB 332 IRSRAQNNNTLKQIYIKLREOFENKTIIVFTSSGGDEIVMHSFNGGGEFYCNSTOLFNS 391
QY 296 TWANNTEGSNNTEGNTITLPCRKOIINMOEYKAMVAPPIRGDIRCSSNITGLLTR 354
DB 392 TW-NTSGSNNTEGNTITLPCRKOIINMOEYKAMVAPPIRGDIRCSSNITGLLTR 450
QY 355 DGIN-ENGTEIFRPGGGDMRDMRSELKYKVKVKEIPLGVAPTCKRRVVOREKRAVGIG 413
DB 451 DGEDTNTTEIFRPGGGDMRDMRSELKYKVKVKEIPLGVAPTCKRRVVOREKRAVGIG 510
QY 414 GAVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOQNNLLRAIEAOQRLQLTWGIKQL 473
DB 511 GALFLGFLGAAGSTMGASMTLTVQARQLLSGIVQOQNNLLRAIEAOQRLQLTWGIKQL 570
QY 474 LQARVLAVERYLGDQQLLGITWGSGLICTTAVPWNASWSKSLDRINNNMTWMEWEREID 533
DB 571 LQARVLAVERYLKDQQLLGITWGSGLICTTAVPWNASWSKSLDRINNNMTWMEWEREID 630

```


Fri Dec 12 15:07:13 2003

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 12:28:04 ; Search time 13.1662 Seconds
(without alignments)
1860.670 Million cell updates/sec

Title: US-10-032-162-15

Perfect score: 3129
Sequence: 1 VEKLVTVYGVVWKEAT.....ELDKWASLWNPDTNWLWY 579Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents, AA:*

1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/ECTUS_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfile1.pep:*Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2800.5	89.5	855	1 US-08-022-835-6	Sequence 6, Appl
2	2800.5	89.5	855	1 US-08-388-809-6	Sequence 6, Appl
3	2800.5	89.5	855	1 US-08-647-714-6	Sequence 6, Appl
4	2776.5	88.7	850	2 US-08-448-603A-28	Sequence 28, Appl
5	2776.5	88.7	850	2 US-09-134-075-28	Sequence 28, Appl
6	2776.5	88.7	850	4 US-09-492-739-28	Sequence 28, Appl
7	2754.5	88.0	855	3 US-07-956-483-14	Sequence 14, Appl
8	2725	87.1	857	2 US-08-448-603A-30	Sequence 30, Appl
9	2725	87.1	857	4 US-09-492-739-30	Sequence 30, Appl
10	2724.5	87.1	861	1 US-08-127-499A-14	Sequence 14, Appl
11	2724.5	87.1	861	1 US-08-482-867-14	Sequence 14, Appl
12	2724.5	87.1	861	1 US-07-956-483-10	Sequence 10, Appl
13	2724.5	87.1	861	3 US-08-472-240A-7	Sequence 7, Appl
14	2724.5	87.1	861	3 US-08-472-240A-1	Sequence 1, Appl
15	2724.5	87.1	861	4 US-08-817-441-103	Sequence 103, App
16	2724.5	86.9	856	4 US-09-124-900-9	Sequence 9, Appl
17	2719	86.9	856	3 US-08-463-210-11	Sequence 11, Appl
18	2719	86.9	856	3 US-08-472-240A-5	Sequence 5, Appl
19	2716.5	86.8	861	3 US-07-956-483-16	Sequence 16, Appl
20	2712	86.7	826	1 US-08-487-657-2	Sequence 2, Appl
21	2712	86.7	826	2 US-08-487-657-2	Sequence 2, Appl
22	2707.5	86.5	880	4 US-09-309-572-23	Sequence 23, Appl
23	2707.5	86.5	880	2 US-08-788-815-7	Sequence 7, Appl
24	2707.5	86.5	880	3 US-09-157-963-7	Sequence 7, Appl
25	2706	86.5	855	3 US-07-956-483-15	Sequence 15, Appl
26	2706	86.4	856	2 US-07-916-098A-2	Sequence 2, Appl
27	2702	86.4	856	2 US-07-916-098A-2	Sequence 2, Appl

28	2695.5	86.1	839	3 US-08-472-240A-10	Sequence 10, Appl
29	2694	86.1	856	4 US-09-337-387-11	Sequence 11, Appl
30	2690	86.0	856	1 US-08-022-835-2	Sequence 2, Appl
31	2690	86.0	856	1 US-08-388-809-2	Sequence 2, Appl
32	2690	86.0	856	2 US-08-647-714-2	Sequence 2, Appl
33	2690	86.0	856	3 US-07-956-483-11	Sequence 11, Appl
34	2686	85.8	857	1 US-08-022-835-4	Sequence 4, Appl
35	2686	85.8	857	1 US-08-388-809-4	Sequence 4, Appl
36	2686	85.8	857	2 US-08-647-714-4	Sequence 4, Appl
37	2672.5	85.4	726	4 US-09-337-387-3	Sequence 3, Appl
38	2670	85.3	887	3 US-08-472-240A-6	Sequence 6, Appl
39	2663	85.1	865	3 US-07-956-483-13	Sequence 13, Appl
40	2654	84.8	856	1 US-08-375-100-1	Sequence 1, Appl
41	2651.5	84.7	759	4 US-09-337-387-12	Sequence 12, Appl
42	2648.5	84.6	657	4 US-09-256-194-2	Sequence 2, Appl
43	2639	84.3	887	3 US-08-472-240A-2	Sequence 2, Appl
44	2626	83.9	887	3 US-08-472-240A-4	Sequence 4, Appl
45	2620.5	83.7	665	2 US-08-448-603A-32	Sequence 32, Appl

ALIGNMENTS

RESULT 1
US-08-022-835-6
; Sequence 6, Application US/08022835
; Patent No. 5420030
; GENERAL INFORMATION:
; APPLICANT: Reitz Jr., Marvin S.
; APPLICANT: Franchini, Genoveffa
; APPLICANT: Markham, Phillip D.
; APPLICANT: Gallo, Robert C.
; APPLICANT: Lorr, Franco C.
; APPLICANT: Popovic, Mikulas
; APPLICANT: Garner, Suzanne
; TITLE OF INVENTION: MOLECULAR CLONES OF HIV-1 AND USES
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN, DABRY & CUSHMAN
STREET: Eleventh floor, 1615 L. Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036-5601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/022,835
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/599,491
FILING DATE: 17-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Scott, Watson T.
REGISTRATION NUMBER: 26,581
REFERENCE/DOCKET NUMBER: WTS/5683/78507/SRL
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 855 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-022-835-6

Query Match 89.5%; Score 2800.5; DB 1; Length 855;
 Best Local Similarity 81.2%; Pred. No. 1.5e-233;
 Matches 527; Conservative 20; Mismatches 31; Indels 71; Gaps 2;

QY 2 EKLWTVYGVVWKEATTTTLFCASDAKAYDEVHANNWATHACVPTDNPQEVLENVTE 61
 DB 31 EKLWTVYGVVWKEATTTTLFCASDRKAYDEVHANNWATHACVPTDNPQEVLENVTE 90
 QY 62 HNNMKNMVEQOMEDIISLMDOSLKPCVKLTPLC----- 96
 DB 91 HNNMKNMVEQOMEDIISLMDOSLKPCVKLTPLC----- 150
 QY 97 --GAG-----CPTSITVITQACP 110
 DB 151 MVGGEMKNGCSFNITTNIRGVQKEYALFYKLDIAPIDNNSNNRYRLSCNTSVITQACP 210
 QY 111 KISFEPPIHCAAPGPAALKNDKTENGKPCPKAVSTVOCTHGRVSVTQLLNGSLA 170
 DB 211 KVSFEPPIHCAAPGPAALKNDKCKKFKNGKPCPTVSTVOCTHGRVSVTQLLNGSLA 270
 171 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGPRAFYTTGELIGDIR 230
 DB 271 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGPRAFYTTGELIGDIR 330
 QY 231 QAHGNISRAKNDTLKOIVIKLRQEPENKTIIVFNHSSGDEIVMHSFNGCGEFPYCNS 290
 DB 331 QAHGNISRAKNDTLKOIVIKLRQEPENKTIIVFNHSSGDEIVMHSFNGCGEFPYCNS 390
 QY 291 QLFNSTWNNTEGSSNNTGNTTLPFCRIKOIINMVEGKAMVAPPIRGQIRCSNITGL 350
 DB 391 QLFNSTWNNTEGSSNNTGNTTLPFCRIKOIINMVEGKAMVAPPIRGQIRCSNITGL 450
 QY 351 LLTRDGGINGETLFRPGGDMRDNRSELYKKVYKIEPLGVAPTKCKRRVQREKRA 410
 DB 451 LLTRDGGEDKTEVFRPGGDMRDNRSELYKKVYKIEPLGVAPTKCKRRVQREKRA 510
 QY 411 VGIGAVFLGFLGAGSTWGAASMTLTVQARLLLSGIVQOONNLRAIEAQOQMLQITVWG 470
 DB 511 VGIGAVFLGFLGAGSTWGAASMTLTVQARLLLSGIVQOONNLRAIEAQOQMLQITVWG 570
 QY 471 IKQIARVLAVERYLGDQOLIGTGCCKLICTTAVPNNASWSKSLDRINNNTTWMEWE 530
 DB 571 IKQIARVLAVERYLGDQOLIGTGCCKLICTTAVPNNASWSKSLDRINNNTTWMEWE 630
 QY 531 REINNTSEIYTLIEESONOEKNEOELELDKASLNNWPDITKMLMY 579
 DB 631 REINNTSEIYTLIEESONOEKNEOELELDKASLNNWPDITKMLMY 679

US-08-388-809-6
 Sequence 6, Application US/08388809
 Patent No. 5576000
 GENERAL INFORMATION:
 APPLICANT: REITZ, JR., MARVIN S.; FRANCHINI,
 APPLICANT: GENOVEFA; MARKHAM, PHILIP D.; GALLO, ROBERT
 APPLICANT: C.; LORI, FRANCO C.; POPOVIC, MIKULAS; AND
 APPLICANT: CARTNER, SUZANNE
 TITLE OF INVENTION: MOLECULAR CLONES OF HIV-1
 TITLE OF INVENTION: AND USES THEREOF
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MORGAN & PINNEGAN
 STREET: 345 PARK AVENUE
 CITY: NEW YORK
 STATE: NEW YORK
 COUNTRY: USA
 ZIP: 10154
 COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY DISK, 3.5"
 COMPUTER: IBM PC COMPATIBLE
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WORDPERFECT 5.1

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/388, 809
 FILING DATE: 15-FEB-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/022,835
 FILING DATE: 25-FEB-1993
 APPLICATION NUMBER: US 07/599,491
 FILING DATE: 17-OCT-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: LESLIE A. SERUNIAN
 REGISTRATION NUMBER: 35,353
 REFERENCE/DOCKET NUMBER: 2026-4092052
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 758-4800
 TELEFAX: (212) 751-6849
 TELEX: 421792
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 855 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-388-809-6

Query Match 89.5%; Score 2800.5; DB 1; Length 855;
 Best Local Similarity 81.2%; Pred. No. 1.5e-233;
 Matches 527; Conservative 20; Mismatches 31; Indels 71; Gaps 2;

QY 2 EKLWTVYGVVWKEATTTTLFCASDAKAYDEVHANNWATHACVPTDNPQEVLENVTE 61
 DB 31 EKLWTVYGVVWKEATTTTLFCASDRKAYDEVHANNWATHACVPTDNPQEVLENVTE 90
 QY 62 HNNMKNMVEQOMEDIISLMDOSLKPCVKLTPLC----- 96
 DB 91 HNNMKNMVEQOMEDIISLMDOSLKPCVKLTPLC----- 150
 QY 97 --GAG-----CPTSITVITQACP 110
 DB 151 MVGGEMKNGCSFNITTNIRGVQKEYALFYKLDIAPIDNNSNNRYRLSCNTSVITQACP 210
 QY 111 KISFEPPIHCAAPGPAALKNDKTENGKPCPKAVSTVOCTHGRVSVTQLLNGSLA 170
 DB 211 KVSFEPPIHCAAPGPAALKNDKCKKFKNGKPCPTVSTVOCTHGRVSVTQLLNGSLA 270
 171 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGPRAFYTTGELIGDIR 230
 DB 271 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGPRAFYTTGELIGDIR 330
 QY 231 QAHGNISRAKNDTLKOIVIKLRQEPENKTIIVFNHSSGDEIVMHSFNGCGEFPYCNS 290
 DB 331 QAHGNISRAKNDTLKOIVIKLRQEPENKTIIVFNHSSGDEIVMHSFNGCGEFPYCNS 390
 QY 291 QLFNSTWNNTEGSSNNTGNTTLPFCRIKOIINMVEGKAMVAPPIRGQIRCSNITGL 350
 DB 391 QLFNSTWNNTEGSSNNTGNTTLPFCRIKOIINMVEGKAMVAPPIRGQIRCSNITGL 450
 QY 351 LLTRDGGINGETLFRPGGDMRDNRSELYKKVYKIEPLGVAPTKCKRRVQREKRA 410
 DB 451 LLTRDGGEDKTEVFRPGGDMRDNRSELYKKVYKIEPLGVAPTKCKRRVQREKRA 510
 QY 411 VGIGAVFLGFLGAGSTWGAASMTLTVQARLLLSGIVQOONNLRAIEAQOQMLQITVWG 470
 DB 511 VGIGAVFLGFLGAGSTWGAASMTLTVQARLLLSGIVQOONNLRAIEAQOQMLQITVWG 570
 QY 471 IKQIARVLAVERYLGDQOLIGTGCCKLICTTAVPNNASWSKSLDRINNNTTWMEWE 530
 DB 571 IKQIARVLAVERYLGDQOLIGTGCCKLICTTAVPNNASWSKSLDRINNNTTWMEWE 630
 QY 531 REINNTSEIYTLIEESONOEKNEOELELDKASLNNWPDITKMLMY 579
 DB 631 REINNTSEIYTLIEESONOEKNEOELELDKASLNNWPDITKMLMY 679

RESULT 3
US-08-647-714-6
Sequence 6, Application US/08647714
Patent No. 5869313
GENERAL INFORMATION:
APPLICANT: REITZ, JR., MARVIN S.; FRANCHINI,
APPLICANT: GENOVEFA; MARKHAM, PHILIP D. GALLO, ROBERT
APPLICANT: C.; LORI, FRANCO C.; POPOVIC, MIKULAS; AND
APPLICANT: GARTNER, SUZANNE
TITLE OF INVENTION: MOLECULAR CLONES OF HIV-1
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & PINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK, 3.5"
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/647,714
FILING DATE:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/08/388,809
FILING DATE: 15-FEB-1995
APPLICATION NUMBER: US 08/022,835
FILING DATE: 25-FEB-1993
APPLICATION NUMBER: US 07/599,491
FILING DATE: 17-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: LESLIE A. SERUNIAN
REGISTRATION NUMBER: 35,353
REFERENCE/DOCKET NUMBER: 2026-4092US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 855 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
DB-647-714-6
Query Match 89.5%; Score 2800.5; DB 2; Length 855;
Best Local Similarity 81.2%; Pred. No. 1.5e-233;
Matches 527; Conservative 20; Mismatches 31; Indels 71; Gaps 2;
QY 2 EKLWTVVYGVVPMKAEATTTTLCASDAKAYDDEVANWATHACVPTDNPQEVLENTYE 61
DB 31 EKLWTVVYGVVPMKAEATTTTLCASDAKAYDDEVANWATHACVPTDNPQEVLENTYE 90
QY 62 HPMKAKNNWQEMODIISLMDQSLKPCYKLTPLC----- 96
DB 91 NFMNKNMNMVQEMEDIISLMDQSLKPCYKLTPLC----- 150
QY 97 --GAG-----CDTSVTOACP 110
DB 151 MVGGEKMKSCNITNTNIGKVKQKAYALFYKLDIAPIDNNSNNRYRLISCNTSVITQACP 210
QY 111 KISEPPIHYCAPAGAILKCNKDTFNGKGPCKNVSTVOCTHGRIPVSTOLLNGSLA 170
DB 211 KVSFEPIDHYCAPAGAILKCKDKKFKNGKPCCTVSTVOCTHGRIPVSTOLLNGSLA 270
QY 171 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSHIGGRAPFTTGEIIGDIR 230
DB 271 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSHIGGRAPFTTGEIIGDIR 330

QY 231 QAHCHISRAKNDITLKOIVIKLREQENKTIIVFNHSGGDEIVNHSFNCGEFFYCNS 290
DB 331 QAHCHISRAKNDITLKOIVIKLREQENKTIIVFNHSGGDEIVNHSFNCGEFFYCNS 390
QY 291 QFNSTNNNTNTEGNNTEGNTITTLPCRIKOIIMNMQEVKAMYPAPIRGOIRCSNTGL 350
DB 391 QFNSTNNNTNTEGNNTEGNTITTLPCRIKOIIMNMQEVKAMYPAPIRGOIRCSNTGL 450
QY 351 LITRGGINGENGTETFRGGGDMRNMRSELYKVVVIEPLGVAPTCKRVRVQREKA 410
DB 451 LITRGGEDNKTVEFRGGGDMRNMRSELYKVVVIEPLGVAPTAKRVRVQREKA 510
QY 411 VIGAVFLGFLGAGSTWGAASMTLTVQARLLLSGIVOQNNLLRAIEAQRMLOLTWYG 470
DB 511 VIGAVFLGFLGAGSTWGAASMTLTVQARLLLSGIVOQNNLLRAIEAQRMLOLTWYG 570
QY 471 IKOLQARVLAVERVYIGDOOLLGIWCGSGKLICTTAVPNNASWSNKSRLDRINNNMTWME 530
DB 571 IKOLQARVLAVERVYLRDQOLLGIWCGSGKLICTTAVPNNASWSNKSRLDRINNNMTWME 630
QY 531 REINDYTSIYTLIESONQOEKNEQLLELDKXASLNNWFDITWMLY 579
DB 631 REINDYTSIYTLIESONQOEKNEQLLELDKXASLNNWFDITWMLY 679
RESULT 4
US-08-448-603A-28
Sequence 28, Application US/08448603A
Patent No. 5864027
GENERAL INFORMATION:
APPLICANT: Berman, Phillip W.
APPLICANT: Nakamura, Gerald R.
TITLE OF INVENTION: HIV Envelope Polypeptides
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: McCutchen, Doyle, Brown & Enersen, LLP
STREET: 3 Embarcadero Center
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/448,603A
FILING DATE: 07-JUN-1994
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/072,833
FILING DATE: 07-JUN-93
ATTORNEY/AGENT INFORMATION:
NAME: Haliday, Emily
REGISTRATION NUMBER: 38903
REFERENCE/DOCKET NUMBER: 14918-704
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-393-2000
TELEFAX: 415-393-2286
TELEX:
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 850 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-448-603A-28
Query Match 88.7%; Score 2776.5; DB 2; Length 850;
Best Local Similarity 81.5%; Pred. No. 1.7e-231;
Matches 528; Conservative 19; Mismatches 28; Indels 73; Gaps 5;

```

QY 2 EKLWVTVYGVVWPKKATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNTE 61
DB 31 EKLWVTVYGVVWPKKATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNTE 90
QY 62 HFNWKNMWEQMOEDIIISLMDQSLKPCVKLTPLCGA----- 98
DB 91 NFNWKNMWEQMOEDIIISLMDQSLKPCVKLTPLCGA----- 150
QY 99 -----GDDTSVITQACPKISFEP 116
DB 151 EIKKCSFNWTTISIRDKKKNKNEVALFYKLDVVIDNDNTSYRLISCNSTVITQACPKISFEP 210
QY 117 IPIHYCAPAGPAILKCKNDKTFNGKPCKNVSTVOCTHGIRPVSTOLLNGSLAEBEVVI 176
DB 211 IPIHYCAPAGPAILKCKNDKTFNGKPCKNVSTVOCTHGIRPVSTOLLNGSLAEBEVVI 270
QY 177 RSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGPRAFYTTGEIIGDIRQAHONI 236
DB 271 RSNFSDNAKTIIVOLNBSVEINCTRPNNNTRRSIHIGPRAFYATGEIIGDIRQAHONI 330
QY 237 SRANKNDLKOIYKLEQENKTIIVFNHSSGDEPEIVMHSFNGGGEFFYCNSTOLFNST 296
DB 331 SSTRKNNLTKOIVTKLREHF-NKTIIVFNHSSGDEPEIVMHSFNGGGEFFYCNSTOLFNST 389
QY 297 WN-----NNTGSGNNTGNTITLPCRKOIINMNOEVGKAMVAPPIRGQIRCSNITGL 351
DB 390 WNYTYTNNTGSDNT-GRNITLQCRKOIINMNOEVGKAMVAPPIRGQIRCSNITGL 448
QY 352 LTRDGINENGTETFRPGGDMRDMRSELYKYKVKIEPLGVAFTKCRKRVVOREKRAV 411
DB 449 LTRDGG-NNSETEIFRPGGDMRDMRSELYKYKVKIEPLGVAFTKCRKRVVOREKRAV 507
QY 412 GIGAVPLGFLGAAGSTWGAASMTLTVQARLLSGIVQOONNLLRAIEAOHLLQLTWGI 471
DB 508 GIGAVPLGFLGAAGSTWGAASMTLTVQARLLSGIVQOONNLLRAIEAOHLLQLTWGI 567
QY 472 KOLARVLAVERYLKDDQDLGIMGCGSKLICTTAVPNAWSNKSILDRINNMTMWER 531
DB 568 KOLARVLAVERYLKDDQDLGIMGCGSKLICTTAVPNAWSNKSILDRINNMTMWER 627
QY 532 EIDNVTSEIYTLIEESQNOEKNEQELLEDKMASLNNWFDITWMLWY 579
DB 628 EIDNVTSEIYTLIEESQNOEKNEQELLEDKMASLNNWFDITWMLWY 675

```

RESULT 5
US-09-134-075-28
Sequence 28, Application US/09134075
Patent No. 6042836
GENERAL INFORMATION:
APPLICANT: Berman, Phillip W.
APPLICANT: Nakamura, Gerald R.
TITLE OF INVENTION: HIV Envelope Polypeptides
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: McCutchen, Doyle, Brown & Enersen, LLP
STREET: 3 Embarcadero Center
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/134,075
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/448,603

```

FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Haliday, Emily
REGISTRATION NUMBER: 38903
REFERENCE/DOCKET NUMBER: 14918-704
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-393-2000
TELEFAX: 415-393-2286
TELEX:
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 850 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-134-075-28
Query Match 88.7%; Score 2776.5; DB 3; Length 850;
Best Local Similarity 81.5%; Pred. No. 1.7e-231;
Matches 528; Conservative 19; Mismatches 28; Indels 73; Gaps 5;

```

```

QY 2 EKLWVTVYGVVWPKKATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNTE 61
DB 31 EKLWVTVYGVVWPKKATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNTE 90
QY 62 HFNWKNMWEQMOEDIIISLMDQSLKPCVKLTPLCGA----- 98
DB 91 NFNWKNMWEQMOEDIIISLMDQSLKPCVKLTPLCGA----- 150
QY 99 -----GDDTSVITQACPKISFEP 116
DB 151 EIKKCSFNWTTISIRDKKKNKNEVALFYKLDVVIDNDNTSYRLISCNSTVITQACPKISFEP 210
QY 117 IPIHYCAPAGPAILKCKNDKTFNGKPCKNVSTVOCTHGIRPVSTOLLNGSLAEBEVVI 176
DB 211 IPIHYCAPAGPAILKCKNDKTFNGKPCKNVSTVOCTHGIRPVSTOLLNGSLAEBEVVI 270
QY 177 RSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGPRAFYTTGEIIGDIRQAHONI 236
DB 271 RSNFSDNAKTIIVOLNBSVEINCTRPNNNTRRSIHIGPRAFYATGEIIGDIRQAHONI 330
QY 237 SRANKNDLKOIYKLEQENKTIIVFNHSSGDEPEIVMHSFNGGGEFFYCNSTOLFNST 296
DB 331 SSTRKNNLTKOIVTKLREHF-NKTIIVFNHSSGDEPEIVMHSFNGGGEFFYCNSTOLFNST 389
QY 297 WN-----NNTGSGNNTGNTITLPCRKOIINMNOEVGKAMVAPPIRGQIRCSNITGL 351
DB 390 WNYTYTNNTGSDNT-GRNITLQCRKOIINMNOEVGKAMVAPPIRGQIRCSNITGL 448
QY 449 LTRDGINENGTETFRPGGDMRDMRSELYKYKVKIEPLGVAFTKCRKRVVOREKRAV 411
DB 508 GIGAVPLGFLGAAGSTWGAASMTLTVQARLLSGIVQOONNLLRAIEAOHLLQLTWGI 567
QY 472 KOLARVLAVERYLKDDQDLGIMGCGSKLICTTAVPNAWSNKSILDRINNMTMWER 531
DB 568 KOLARVLAVERYLKDDQDLGIMGCGSKLICTTAVPNAWSNKSILDRINNMTMWER 627
QY 532 EIDNVTSEIYTLIEESQNOEKNEQELLEDKMASLNNWFDITWMLWY 579
DB 628 EIDNVTSEIYTLIEESQNOEKNEQELLEDKMASLNNWFDITWMLWY 675

```

RESULT 6
US-09-492-739-28
Sequence 28, Application US/09492739
Patent No. 6331404
GENERAL INFORMATION:
APPLICANT: Berman, Phillip W.
APPLICANT: Nakamura, Gerald R.

TITLE OF INVENTION: HIV Envelope Polypeptides
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: McCutchen, Doyle, Brown & Eneersen, LLP
STREET: 3 Embarcadero Center
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/492,739
FILING DATE: 27-Jan-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/134,075
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Haliday, Emily
REGISTRATION NUMBER: 38903
REFERENCE/DOCKET NUMBER: 14918-704
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-393-2000
TELEFAX: 415-393-2286
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 850 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 28:
US-09-492-739-28

Query Match 88.7% Score 2776.5; DB 4; Length 850;
Best Local Similarity 81.5%; Pred. No. 1.7e-231;
Matches 528; Conservative 19; Mismatches 28; Indels 73; Gaps 5;

QY 2 EKLAVTVVYGVVPMKEATTTLLFCASDAKAYDTEVNNVWATTHACVPTDPNPOEVLNTE 61
DB 31 EKLAVTVVYGVVPMKEATTTLLFCASDAKAYDTEVNNVWATTHACVPTDPNPOEVLNTE 90
QY 62 HNNMKNMNVQEMODIISLMDQSLKPCVKLTPLCGA----- 98
DB 91 NNNMKNMNVQEMODIISLMDQSLKPCVKLTPLCGA----- 150
QY 99 -----GCDTSVITQACPKISFEP 116
DB 151 EIKNCSFNTTISIRDMKNKNEVALFYKLDVPIIDNDNTSYRLISCTNSVITQACPKISFEP 210
QY 117 IPIHYCAPAGFAILKCNKDTFNGKPCKNVSTVQCTHGIRPVSTQLLNGSLAESEVYI 176
DB 211 IPIHYCAPAGFAILKCRDKKFGTGTCTVSTVQCTHGIRPVSTQLLNGSLAESEVYI 270
QY 177 RSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSHIGGAPFYTTGIIIGDIRAHQNI 236
DB 271 RSNANFSDNAKTIIVOLKESVEINCTRPNNNTRKSHIGGAPFYTTGIIIGDIRAHQNI 330
QY 237 SAKKNDTLKOIVIKRBOFENKTIIVNHSNGGDEPIVNSFCGSEFPYCSTQLFNST 296
DB 331 SSKNNNTLKOIVIKRBOFENKTIIVNHSNGGDEPIVNSFCGSEFPYCSTQLFNST 389
QY 297 WN-----NNTGSSNNTGNTITLPCRKOIINMMQEVGKAMYAPIRGOIRGSSNTTGL 351
DB 399 WNYTNTMNTGSSNTI-GRNITLQCRKIKOIINMMQEVGKAMYAPIRGOIRGSSNTTGL 448
QY 352 LTRDGINENGTETIRPGGGMRRDNNRSLYKVKVKEIPLGVAFTKCRRAVQREKRAV 411
DB 449 LTRDGG-NNSETEIFRPGGGMRRDNNRSELKYKVKVKEIPLGVAFTKCRRAVQREKRAV 507

QY 412 GIGAVFLGFLGAGSTMGASMTLTVOARLLLSGIVQOQNNLLRAIEAQORMLQITWGI 471
DB 508 GIGAVFLGFLGAGSTMGASMTLTVOARLLLSGIVQOQNNLLRAIEAQORMLQITWGI 567
QY 472 KOLQARVLAVERYLADQDLGIGWCSGKLTCTVAPNNASNSKSLDIINNNMTMEWER 531
DB 568 KOLQARVLAVERYLADQDLGIGWCSGKLTCTVAPNNASNSKSLDIINNNMTMEWER 627
QY 532 EIDNTSRIYTLIERSONQOEKNEBELLEDKMASLWMPDITWLMY 579
DB 628 EIDNTSRIYTLIERSONQOEKNEBELLEDKMASLWMPDITWLMY 675

RESULT 7
US-07-956-483-14
Sequence 14, Application US/07956483
Patent No. 6261799
GENERAL INFORMATION:
APPLICANT: KIENY, Marie-Paule
TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED
TITLE OF INVENTION: gp160 VARIANT
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/956,483
FILING DATE: 31-DEC-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 92/19742
FILING DATE: 12-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 05392
FILING DATE: 02-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Crane-Feury, Sharon E
REGISTRATION NUMBER: 36,113
REFERENCE/DOCKET NUMBER: 017753-005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 855 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-07-956-483-14

Query Match 88.0% Score 2754.5; DB 3; Length 855;
Best Local Similarity 80.1%; Pred. No. 1.4e-229;
Matches 523; Conservative 26; Mismatches 27; Indels 77; Gaps 6;

QY 2 EKLAVTVVYGVVPMKEATTTLLFCASDAKAYDTEVNNVWATTHACVPTDPNPOEVLNTE 61
DB 31 EKLAVTVVYGVVPMKEATTTLLFCASDAKAYDTEVNNVWATTHACVPTDPNPOEVLNTE 90
QY 62 HNNMKNMNVQEMODIISLMDQSLKPCVKLTPLCGA----- 98
DB 91 NNNMKNMNVQEMODIISLMDQSLKPCVKLTPLCGA----- 150

```

QY 99 -----GCDTSVITQACPKIS 113
DB 151 KMEGEMTNCNSFNITTSIRSKVQKEVALFYKLDVVPIDNTSYTLINCTSVITQACPKIS 210
QY 114 FEPPIPIHYCAPAGAILLKNDXTNGKGPCKNVSTVQCTHGIRPVSTOLLNLSLAEE 173
DB 211 FEPPIPIHYCA-RWFAIILNCKNKKFNGTGPCTNVSTVQCTHGIRPVSTOLLNLSLAEE 269
QY 174 VVISDSNFTNNAKITIVOLKESVEINCTRPNNNRKSIHIGRGAFTTGEIIGDIOAH 233
DB 270 VVLSSENFIDAKITIVOLKESVEINCTRPNNNRKSIHIGRGAFTTGEIIGDIOAH 329
QY 234 CNISRAKMNIDLKQIVIKLREOFENKTIIVFNHSSGDEPEIVMHSFNGCGEFCYCNSTOLF 293
DB 330 CNISRAKMNNTLKQIVIKLREOFENKTIIVFNHSSGDEPEIVMHSFNGCGEFCYCNSTOLF 389
QY 294 NSTNNNTNTEGNTN-ITTLPCRIKQIIMMOEYKAMAPPIRGQIRCSSITGILL 352
DB 390 SST-NGTEGNNNGDITLPCRIKEIIMMOEYKAMAPPIRGQIRCSSITGILL 448
QY 353 TROGGINENG-----TEIFRPGGDMRDNMRSELYKYVVKIEPLGVAPTCKRKYORE 407
DB 449 TRDGSNKSNGSKNENTEIFRPGGDMRDNMRSELYKYVVKIEPLGVAPTCKRKYORE 508
QY 408 KRAVG-IGAVELGFLGAAGSTMGASMTLTVQARLLLSGIVQOQNLLRAIEAOQRLQL 466
DB 509 KRAVGTIGAVELGFLGAAGSTMGASMTLTVQARLLLSGIVQOQNLLRAIEAOQRLQL 568
QY 467 TWGIKQOLARVLAVERYLGDQOLGIGWCSGKLICTTAVPWNASWSNKSIDRIWNNMTW 526
DB 569 TWGIKQOLARVLAVERYLGDQOLGIGWCSGKLICTTAVPWNASWSNKSIDRIWNNMTW 628
QY 527 MEMREIDNTSEIYTLIESONQOEKNEDELLELDKASLNNWPDITWMLY 579
DB 629 MEMREIDNTSEIYTLIESONQOEKNEDELLELDKASLNNWPDITWMLY 681

```

RESULT 8

```

US-08-448-603A-30
; Sequence 30, Application US/08448603A
; Patent No. 3864027

```

```

; GENERAL INFORMATION:
; APPLICANT: Berman, Phillip W.
; APPLICANT: Nakamura, Gerald R.
; TITLE OF INVENTION: HIV Envelope Polypeptides
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESS: McCutchen, Doyle, Brown & Enersen, LLP
; STREET: 3 Embarcadero Center
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: PASCAL for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448.603A
; FILING DATE: 07-JUN-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/072.833
; FILING DATE: 07-JUN-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Halliday, Emily
; REGISTRATION NUMBER: 38903
; REFERENCE/DOCKET NUMBER: 14918-704
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-393-2000
; TELEFAX: 415-393-2286
; TELEX:

```

```

; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 857 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-448-603A-30

```

```

Query Match 87.1%; Score 2725; DB 2; Length 857;
Best Local Similarity 78.6%; Pred. No. 5e-227;
Matches 512; Conservative 31; Mismatches 34; Indels 74; Gaps 3;

```

```

QY 3 KLVWTVYVGVVPEAKETTLFCASDAKAYDEVNANWATHACVPTDPNPOEVLENTEH 62
DB 32 KLVWTVYVGVVPEAKETTLFCASDAKAYDEVNANWATHACVPTDPNPOEVLENTEH 91
QY 63 FNNKKNMNVQOMEDTISLMDQSLKPCVLTPLC-----GAG----- 99
DB 92 FNNKKNMNVQOMEDTISLMDQSLKPCVLTPLC-----GAG----- 151
QY 100 -----GDTSVITQACPKIS 110
DB 152 IKNCSEFNTTSVADKQKETALRNKLDIVPIDDDDRSTNSTNYRLISCTSVITQACPKIS 211
QY 111 KISEPPIPIHYCAPAGAILLKNDXTNGKGPCKNVSTVQCTHGIRPVSTOLLNLSLA 170
DB 212 KISEPPIPIHYCAPAGAILLKNDXTNGKGPCKNVSTVQCTHGIRPVSTOLLNLSLA 271
QY 171 EEEVVISDSNFTNNAKITIVOLKESVEINCTRPNNNRKSIHIGRGAFTTGEIIGDIR 230
DB 272 EEEVVISDSNFTNNAKITIVOLKESVEINCTRPNNNRKSIHIGRGAFTTGEIIGDIR 331
QY 231 QAHNISRANKNDLTKQIVIKLREOFENKTIIVFNHSSGDEPEIVMHSFNGCGEFCYCNST 290
DB 332 QAHNISRANKNDLTKQIVIKLREOFENKTIIVFNHSSGDEPEIVMHSFNGCGEFCYCNST 391
QY 392 QLFSTWNTDKVSGNSTEENSTTLPCRIKQIIMMOEYKAMAPPIRGQIRCSSIT 451
DB 452 QLFSTWNTDKVSGNSTEENSTTLPCRIKQIIMMOEYKAMAPPIRGQIRCSSIT 511
QY 511 GLLTTRDGSNKSNGSKNENTEIFRPGGDMRDNMRSELYKYVVKIEPLGVAPTCKRKYORE 568
DB 612 GLLTTRDGSNKSNGSKNENTEIFRPGGDMRDNMRSELYKYVVKIEPLGVAPTCKRKYORE 671
QY 671 RAVGIGAVELGFLGAAGSTMGASMTLTVQARLLLSGIVQOQNLLRAIEAOQRLQL 728
DB 728 RAVGIGAVELGFLGAAGSTMGASMTLTVQARLLLSGIVQOQNLLRAIEAOQRLQL 785
QY 785 WGIKQOLARVLAVERYLGDQOLGIGWCSGKLICTTAVPWNASWSNKSIDRIWNNMTW 842
DB 842 WGIKQOLARVLAVERYLGDQOLGIGWCSGKLICTTAVPWNASWSNKSIDRIWNNMTW 899
QY 899 MEMREIDNTSEIYTLIESONQOEKNEDELLELDKASLNNWPDITWMLY 956
DB 956 MEMREIDNTSEIYTLIESONQOEKNEDELLELDKASLNNWPDITWMLY 1013

```

RESULT 9

```

US-09-134-075-30
; Sequence 30, Application US/09134075
; Patent No. 6042836

```

```

; GENERAL INFORMATION:
; APPLICANT: Berman, Phillip W.
; APPLICANT: Nakamura, Gerald R.
; TITLE OF INVENTION: HIV Envelope Polypeptides
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESS: McCutchen, Doyle, Brown & Enersen, LLP
; STREET: 3 Embarcadero Center
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/09/134,075
  FILING DATE:
  CLASSIFICATION:
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: 08/448,603
  FILING DATE:
  ATTORNEY/AGENT INFORMATION:
    NAME: Haliday, Emily
    REGISTRATION NUMBER: 38903
    REFERENCE/DOCKET NUMBER: 14918-704
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 415-393-2000
    TELEFAX: 415-393-2286
  TELEX:
  INFORMATION FOR SEQ ID NO: 30:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 857 amino acids
    TYPE: amino acid
    STRANDEDNESS: single
    TOPOLOGY: linear
  US-09-134-075-30

```

Query Match 87.1%; Score 2725; DB 3; Length 857;
Best Local Similarity 78.6%; Pred. No. 5e-227;

Matches 512; Conservative 31; Mismatches 34; Indels 74; Gaps 3;

```

QY 3 KLMVTVVYGVVPMKATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENVTEH 62
DB 32 KLMVTVVYGVVPMKATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENVTEH 91
QY 63 FNMKNVNEQWQEDIIISLDQSLKPCVLTPLC-----GAG----- 99
DB 92 FNMKNVNEQWQEDIIISLDQSLKPCVLTPLC-----GAG----- 151
QY 100 -----CQTSVITQACP 110
DB 152 IKNSFNITTSVRDKQKETALENKLDIVPIDDDRNSTRNSTNYRLISCNTSVITQACP 211
QY 111 KISFEPIDHYCAPAGFALIKCNDKTFNGKPCCKNVSTVQCTHGIRPVVSTOLLNGSLA 170
DB 212 KVSFEPIDHYCAPAGFALIKCNDKTFNGKPCCKNVSTVQCTHGIRPVVSTOLLNGSLA 271
QY 171 EBEVVISRDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGRGAFTTGEIIGDIR 230
DB 272 EBEVVISRDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGRGAFTTGEIIGDIR 331
QY 231 QAHCHISRAKANDTLKQIYIKLREOFENKTIIVNHSGGDPRIVMSFNGSGEFPYCNST 290
DB 332 QAHCHISRAKANDTLKQIYIKLREOFENKTIIVNHSGGDPRIVMSFNGSGEFPYCNST 391
QY 291 QLFNSTNNNN--TBSNNTSGNTITLPCGIKQIIMNMOVGKAMAPPRGQIRCSSNT 348
DB 392 QLFNSTNNNN--TBSNNTSGNTITLPCGIKQIIMNMOVGKAMAPPRGQIRCSSNT 451
QY 349 GILLTRDGINENGTEIFRPGGDMRDNRSELKYKVKIPLGLVAPFKCARVVOREK 408
DB 452 GILLTRDGINENGTEIFRPGGDMRDNRSELKYKVKIPLGLVAPFKCARVVOREK 511
QY 409 RAVGIGAVFLGLGAAGSTMGASMTLVQARLLSGIVQOQNNILRAIEAQOQMLQTLV 468
DB 512 RAVGIGAVFLGLGAAGSTMGASMTLVQARLLSGIVQOQNNILRAIEAQOQMLQTLV 571
QY 469 WGIKLOLQVLAVERVYIGDQQLIGTIGSGSKLICTAVPWNASWNSKSLDRITNNMTME 528
DB 572 WGIKLOLQVLAVERVYIGDQQLIGTIGSGSKLICTAVPWNASWNSKSLDRITNNMTME 631
QY 529 WEREIDNTSEIYTLIEESQNOQEKNEOBLLELDKMASLMMNFDITNMLMY 579

```

DB 632 WEREIDNTSEIYTLIEESQNOQEKNEOBLLELDKMASLMMNFDITNMLMY 682

```

RESULT 10
US-09-492-739-30
Sequence 30, Application US/09492739
Patent No. 631404
GENERAL INFORMATION:
  APPLICANT: Berman, Phillip W.
  TITLE OF INVENTION: HIV Envelope Polypeptides
  NUMBER OF SEQUENCES: 33
  CORRESPONDENCE ADDRESS:
    ADDRESS: McCutchen, Doyle, Brown & Eversen, LLP
    STREET: 3 Embarcadero Center
    CITY: San Francisco
    STATE: CA
    COUNTRY: USA
  ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/09/492,739
  FILING DATE: 27-Jan-2000
  CLASSIFICATION: <Unknown>
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: 09/134,075
    FILING DATE: <Unknown>
  ATTORNEY/AGENT INFORMATION:
    NAME: Haliday, Emily
    REGISTRATION NUMBER: 38903
    REFERENCE/DOCKET NUMBER: 14918-704
    TELECOMMUNICATION INFORMATION:
      TELEPHONE: 415-393-2000
      TELEFAX: 415-393-2286
    TELEX: <Unknown>
  INFORMATION FOR SEQ ID NO: 30:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 857 amino acids
    TYPE: amino acid
    STRANDEDNESS: single
    TOPOLOGY: linear
  SEQUENCE DESCRIPTION: SEQ ID NO: 30:
  US-09-492-739-30

```

Query Match 87.1%; Score 2725; DB 4; Length 857;

Best Local Similarity 78.6%; Pred. No. 5e-227;

Matches 512; Conservative 31; Mismatches 34; Indels 74; Gaps 3;

```

QY 3 KLMVTVVYGVVPMKATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENVTEH 62
DB 32 KLMVTVVYGVVPMKATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENVTEH 91
QY 63 FNMKNVNEQWQEDIIISLDQSLKPCVLTPLC-----GAG----- 99
DB 92 FNMKNVNEQWQEDIIISLDQSLKPCVLTPLC-----GAG----- 151
QY 100 -----CQTSVITQACP 110
DB 152 IKNSFNITTSVRDKQKETALENKLDIVPIDDDRNSTRNSTNYRLISCNTSVITQACP 211
QY 111 KISFEPIDHYCAPAGFALIKCNDKTFNGKPCCKNVSTVQCTHGIRPVVSTOLLNGSLA 170
DB 212 KVSFEPIDHYCAPAGFALIKCNDKTFNGKPCCKNVSTVQCTHGIRPVVSTOLLNGSLA 271
QY 171 EBEVVISRDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGRGAFTTGEIIGDIR 230
DB 272 EBEVVISRDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGRGAFTTGEIIGDIR 331

```

QY 231 QAHCHISRAKNDTLKQIVIKLREQFNKTIIVFNHSSGGDEBEIVMHSNCGGEFFYCNS 290
 DB 332 QAHCHISRAKNDTLKQIVIKLREQFNKTIIVFNHSSGGDEBEIVMHSNCGGEFFYCNS 391
 QY 291 QFNSTNNNN--TEGSNNTEGNTITLPCRIKQIINMOEKGKAMVAPPIRQIRCSNIT 348
 DB 392 QFNSTNNNN--TEGSNNTEGNTITLPCRIKQIINMOEKGKAMVAPPIRQIRCSNIT 451
 QY 349 GLLTLRGGIENGTEIFRPGGDMRDWRSELYKVKVKTIEPIGVAPTKRRVVOREK 408
 DB 452 GLLTLRGGIENGTEIFRPGGDMRDWRSELYKVKVKTIEPIGVAPTKRRVVOREK 511
 QY 409 RAVGIGAVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOONNLRAIEAQOEMLOLV 468
 DB 512 RAVGIGAVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOONNLRAIEAQOEMLOLV 571
 QY 469 WGIKOLQARVLAVERYLDQOOLGIGWCSGKLICTTAVPNNASWSNKSIDRIWNNMTWE 528
 DB 572 WGIKOLQARVLAVERYLDQOOLGIGWCSGKLICTTAVPNNASWSNKSIDRIWNNMTWE 631
 QY 529 MERIDNTSEITLIESONQOEKNEBELLELDKMASIMNPFITNNLMY 579
 DB 632 MERIDNTSEITLIESONQOEKNEBELLELDKMASIMNPFITNNLMY 682

RESULT 11

US-08-127-499A-14
 ; Sequence 14, Application US/08127499A
 ; Patent No. 5510264
 ; GENERAL INFORMATION:
 ; APPLICANT: VAN ALSTYNE, Diane
 ; APPLICANT: SHARMA, Lawrence Rajendra
 ; TITLE OF INVENTION: ANTIBODIES WHICH BIND MENINGITIS RELATED
 ; TITLE OF INVENTION: HOMOLOGOUS ANTIGENIC SEQUENCES
 ; NUMBER OF SEQUENCES: 40
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Foley & Lardner
 ; STREET: 3000 K Street, N.W., Suite 500
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20007-5109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/127,499A
 ; FILING DATE: 28-SEP-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: BENT, Stephen A.
 ; REGISTRATION NUMBER: 29,768
 ; REFERENCE/DOCKET NUMBER: 51916/102/INBI
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 672-5300
 ; TELEFAX: (202) 672-5399
 ; TELEX: 904136
 ; INFORMATION FOR SEQ ID NO: 14:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 861 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: unknown
 ; US-08-127-499A-14

Query Match 87.1%; Score 2724.5; DB 1; Length 861;
 Best Local Similarity 79.6%; Pred. No. 5-5e-227;
 Matches 522; Conservative 25; Mismatches 30; Indels 79; Gaps 6;
 QY 2 EKLWTVYVGVVWKEATTTTFCASDAKAYDTEVHNWATACVPTDPNPQEVLEANTTE 61
 DB 32 EKLWTVYVGVVWKEATTTTFCASDAKAYDTEVHNWATACVPTDPNPQEVLEANTTE 91

QY 62 HFNWKKNNWTEQOEDIIISLMDQSLKPCVLTPLCGA----- 98
 DB 92 NFNWKKNNWTEQOEDIIISLMDQSLKPCVLTPLCGA----- 151
 QY 99 -----GCDTSYITQACP 110
 DB 152 MMEKEGKNCSEFNISTIRGKQKQKAYAFYKLDIIPIDNTTSYTLTSCNTSVITQACP 211
 QY 111 KISREPIPIHYCAPAGAILIKNDKTPNGKPCKNVSTVOCTGIRVYSTOLLNSLA 170
 DB 212 KVSSEPIPIHYCAPAGAILIKNDKTPNGKPCKNVSTVOCTGIRVYSTOLLNSLA 271
 QY 171 EEEVVISDFTNNAKTIIVOLKESVEINCTRPNNTRKSIH--GPRAFYTTGEIIGD 228
 DB 272 EEEVVISDFTNNAKTIIVOLKESVEINCTRPNNTRKSIH--GPRAFYTTGEIIGD 330
 QY 229 IQAHCHISRAKNDTLKQIVIKLREQFNKTIIVFNHSSGGDEBEIVMHSNCGGEFFYC 287
 DB 331 IQAHCHISRAKNDTLKQIVIKLREQFNKTIIVFNHSSGGDEBEIVMHSNCGGEFFYC 390
 QY 288 NSTQLFNSSTNN--NTEGSNNTEG--NTITLPCRIKQIINMOEKGKAMVAPPIRQIR 343
 DB 391 NSTQLFNSSTNN--NTEGSNNTEG--NTITLPCRIKQIINMOEKGKAMVAPPIRQIR 450
 QY 344 SSNTTGLLTRDGGIENGTEIFRPGGDMRDWRSELYKVKVKTIEPIGVAPTKRRV 403
 DB 451 SSNTTGLLTRDGGIENGTEIFRPGGDMRDWRSELYKVKVKTIEPIGVAPTKRRV 510
 QY 404 VOREKRAVIGAVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOONNLRAIEAQOEM 463
 DB 511 VOREKRAVIGAVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOONNLRAIEAQOEM 570
 QY 464 LQTVWGIKOLQARVLAVERYLDQOOLGIGWCSGKLICTTAVPNNASWSNKSIDRIWNN 523
 DB 571 LQTVWGIKOLQARVLAVERYLDQOOLGIGWCSGKLICTTAVPNNASWSNKSIDRIWNN 630
 QY 524 MTWMEEREIDNTSEITLIESONQOEKNEBELLELDKMASIMNPFITNNLMY 579
 DB 631 MTWMEEREIDNTSEITLIESONQOEKNEBELLELDKMASIMNPFITNNLMY 686

RESULT 12

US-08-482-847-14
 ; Sequence 14, Application US/08482847
 ; Patent No. 5556757
 ; GENERAL INFORMATION:
 ; APPLICANT: VAN ALSTYNE, Diane
 ; APPLICANT: SHARMA, Lawrence Rajendra
 ; TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPIC SITES FOR
 ; TITLE OF INVENTION: BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR
 ; TITLE OF INVENTION: CNS CARRIER, ANTIBODIES THEREOF, AND USES THEREOF
 ; NUMBER OF SEQUENCES: 40
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Foley & Lardner
 ; STREET: 3000 K Street, N.W., Suite 500
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20007-5109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/482,847
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/127,499
 ; FILING DATE: 28-SEP-1993
 ; ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 51916/104/INBI
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 672-5300
 TELEFAX: (202) 672-5399
 TELER: 904136
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 861 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: unknown

US-08-482-847-14

Query Match 87.1%; Score 2724.5; DB 1; Length 861;
 Best Local Similarity 79.6%; Pred. No. 5.5e-227;
 Matches 522; Conservative 25; Mismatches 30; Indels 79; Gaps 6;

2 EKLMTVYVGVVWKEATTTLLFCASDAKAYDEVHNVMTACVPTDPNPOEVLENTYE 61
 32 EKLMTVYVGVVWKEATTTLLFCASDAKAYDEVHNVMTACVPTDPNPOEVLENTYE 91
 62 HNNMKNVWQEDIIISLMDQSLKPCVKLTPLCGA----- 98
 92 NNNMKNVWQEDIIISLMDQSLKPCVKLTPLCGA----- 151
 99 -----GCDTSVITQACP 110
 152 NMMKEGKINCSFNISTIRGKQKVEAFYKLDIIPIDNDTTSYLTSCNTSVITQACP 211
 111 KISFEPIPIHYCAPAGFALCKNDKTFNGKPCCKNVSTVQCTHGIRPVVSTOLLNGSLA 170
 212 KVSFEPIPIHYCAPAGFALCKNDKTFNGKPCCKNVSTVQCTHGIRPVVSTOLLNGSLA 271
 171 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHI--GPGRAFYTTGEIIGD 228
 272 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHI--GPGRAFYTTGEIIGD 330
 229 IROAHCNISRAKNDTLKQIVIKLREOF--ENKTIYFNHSSGGDEIIVHMSFPCGGGEFFYC 287
 331 MROAHCNISRAKNDTLKQIVIKLREOF--ENKTIYFNHSSGGDEIIVHMSFPCGGGEFFYC 390
 288 NSTOLFNSNTWN---NTEGSNNTGEG--NTITLPCRIRKQIINMGOEVGKAMVAPPIRQIIRC 343
 391 NSTOLFNSNTWN---NTEGSNNTGEG--NTITLPCRIRKQIINMGOEVGKAMVAPPIRQIIRC 450
 344 SSNITGLLLTRDGINENGTEIFRPGGDMRDNRSELYKYKVKIIEPLGVAFTYKAKRRV 403
 451 SSNITGLLLTRDGINENGTEIFRPGGDMRDNRSELYKYKVKIIEPLGVAFTYKAKRRV 510
 404 VOREKRAVIGAVPLGFLGAAGSTWGAASMTLVQARLLSGIVQOONMLRAIEAQOM 463
 511 VOREKRAVIGAVPLGFLGAAGSTWGAASMTLVQARLLSGIVQOONMLRAIEAQOM 570
 464 LQLTWGIQOLARVLAVERYLGDQDLGIGWCSGLICTAVPMWASNSKSLBIMN 523
 571 LQLTWGIQOLARVLAVERYLGDQDLGIGWCSGLICTAVPMWASNSKSLBIMN 630
 524 MTMWEREIDNYTSEIYTLIESSONQOEKNEOELLEDKMASLMMWPDITWMLWY 579
 631 MTMWEREIDNYTSEIYTLIESSONQOEKNEOELLEDKMASLMMWPDITWMLWY 686

RESULT 13
 US-07-956-483-10
 Sequence 10, Application US/07956483
 Patent No. 6261799
 GENERAL INFORMATION:
 APPLICANT: KIENY, Marie-Paule
 TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED
 NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Burns, Doane, Swecker & Mathis
 STREET: P. O. Box 1404
 CITY: Alexandria
 STATE: Virginia
 COUNTRY: United States
 ZIP: 22313-1404
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/956,483
 FILING DATE: 31-DEC-1992
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO 92/19742
 FILING DATE: 12-NOV-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: FR 91 05392
 FILING DATE: 02-MAY-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Crane-Feury, Sharon E
 REGISTRATION NUMBER: 36,113
 REFERENCE/DOCKET NUMBER: 01753-005
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 836-6620
 TELEFAX: (703) 836-2021
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 861 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 US-07-956-483-10

Query Match 87.1%; Score 2724.5; DB 3; Length 861;
 Best Local Similarity 79.6%; Pred. No. 5.5e-227;
 Matches 522; Conservative 25; Mismatches 30; Indels 79; Gaps 6;

2 EKLMTVYVGVVWKEATTTLLFCASDAKAYDEVHNVMTACVPTDPNPOEVLENTYE 61
 32 EKLMTVYVGVVWKEATTTLLFCASDAKAYDEVHNVMTACVPTDPNPOEVLENTYE 91
 62 HNNMKNVWQEDIIISLMDQSLKPCVKLTPLCGA----- 98
 92 NNNMKNVWQEDIIISLMDQSLKPCVKLTPLCGA----- 151
 99 -----GCDTSVITQACP 110
 152 NMMKEGKINCSFNISTIRGKQKVEAFYKLDIIPIDNDTTSYLTSCNTSVITQACP 211
 111 KISFEPIPIHYCAPAGFALCKNDKTFNGKPCCKNVSTVQCTHGIRPVVSTOLLNGSLA 170
 212 KVSFEPIPIHYCAPAGFALCKNDKTFNGKPCCKNVSTVQCTHGIRPVVSTOLLNGSLA 271
 171 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHI--GPGRAFYTTGEIIGD 228
 272 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHI--GPGRAFYTTGEIIGD 330
 229 IROAHCNISRAKNDTLKQIVIKLREOF--ENKTIYFNHSSGGDEIIVHMSFPCGGGEFFYC 287
 331 MROAHCNISRAKNDTLKQIVIKLREOF--ENKTIYFNHSSGGDEIIVHMSFPCGGGEFFYC 390
 288 NSTOLFNSNTWN---NTEGSNNTGEG--NTITLPCRIRKQIINMGOEVGKAMVAPPIRQIIRC 343
 391 NSTOLFNSNTWN---NTEGSNNTGEG--NTITLPCRIRKQIINMGOEVGKAMVAPPIRQIIRC 450
 344 SSNITGLLLTRDGINENGTEIFRPGGDMRDNRSELYKYKVKIIEPLGVAFTYKAKRRV 403
 451 SSNITGLLLTRDGINENGTEIFRPGGDMRDNRSELYKYKVKIIEPLGVAFTYKAKRRV 510

QY 404 VOREKRAVIGAVELFLGAGSTMGASMTLTVQARLLSGIVQOQNNLLRAIEAOQM 463
DB 511 VOREKRAVIGALFLGFLGAGSTMGASMTLTVQARLLSGIVQOQNNLLRAIEAOQH 570
QY 464 LQLTWVGIKQLOARVLAVERYLGDQOLGIGWCSGKLICTTAVPWNASWNSKSLDRITWN 523
DB 571 LQLTWVGIKQLOARVLAVERYLKDQOLGIGWCSGKLICTTAVPWNASWNSKSLDRITWN 630
QY 524 MTWMEEREIDNYSEIYTLIEESONQOEKNEBELLELDKASIMNFDITNMLY 579
DB 631 MTWMEEREIDNYSEIYTLIEESONQOEKNEBELLELDKASIMNFDITNMLY 686

RESULT 14

US-08-472-240A-1

Sequence 1, Application US/08472240A

Patent No. 6284248

GENERAL INFORMATION:

APPLICANT: KIENY, Marie-Paule

TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/472,240A

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/956,483

FILING DATE: 31-DEC-1992

ATTORNEY/AGENT INFORMATION:

NAME: Teskin, Robin L.

REGISTRATION NUMBER: 35,030

REFERENCE/DOCKET NUMBER: 017753-055

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-6620

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 861 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

MOLECULE TYPE: peptide

FEATURE:

NAME/KEY: Protein

LOCATION: 1..854

US-08-472-240A-1

Query Match

Best Local Similarity: 79.1%; Score 2724.5; DB 3; Length 861;

Matches 522; Conservative 25; Mismatches 30; Indels 79; Gaps 6;

QY 2 EKLWTVYVGVPMKEATTLTFCASDAKAYDTEVHNWNTAACTPDNPQSEVLENTYE 61
DB 32 EKLWTVYVGVPMKEATTLTFCASDAKAYDTEVHNWNTAACTPDNPQSEVLENTYE 91
QY 62 HNNMKNMVEQMODIISLMQSLKPCVKLTPLGCA----- 98
DB 92 NNNMKNMVEQMODIISLMQSLKPCVKLTPLGCA----- 151

QY 99 -----GCDTSVITQACP 110
DB 152 MMMEKEIKNCSPNISTSRIGKVQKEYAFYKLDIIPIDNDTTSYLLSCNTSVITQACP 211
QY 111 KISPEPIPIHYCAPAGAILKNDKPTENGKPCKNYSTVCTGIRPVYSTOLLNSLA 170
DB 212 KVSPEPIPIHYCAPAGAILKNDKPTENGKPCKNYSTVCTGIRPVYSTOLLNSLA 271
QY 171 EEEVIRSDNFTNNAKTIIVQLKESVEINCTRNPNTRKSIH--GPRAFYTTGELIGD 228
DB 272 EEEVIRSDNFTNNAKTIIVQLKESVEINCTRNPNTRKSIH--GPRAFYTTGELIGD 330
QY 229 IROAHNCISRAKNWDTLKQIVIKREOF-ENKTIIVNHSAGGDEIYMSFNGGEFFYC 287
DB 331 IROAHNCISRAKNWDTLKQIVIKREOF-ENKTIIVNHSAGGDEIYMSFNGGEFFYC 390
QY 288 NSTOLENSTWNN--NTEGSSNTEG--NTITLPCRIKIIIMMOEYGVKAMAPRGOIRC 343
DB 391 NSTOLENSTWNN--NTEGSSNTEG--NTITLPCRIKIIIMMOEYGVKAMAPRGOIRC 450
QY 344 SSNTGLTLTRDGINNGTEIFRPGGDMRDNRSELYKVKYKIEPLGVAPTKGRV 403
DB 451 SSNTGLTLTRDGINNGTEIFRPGGDMRDNRSELYKVKYKIEPLGVAPTKGRV 510
QY 404 VOREKRAVIGAVELFLGAGSTMGASMTLTVQARLLSGIVQOQNNLLRAIEAOQM 463
DB 511 VOREKRAVIGALFLGFLGAGSTMGASMTLTVQARLLSGIVQOQNNLLRAIEAOQH 570
QY 464 LQLTWVGIKQLOARVLAVERYLGDQOLGIGWCSGKLICTTAVPWNASWNSKSLDRITWN 523
DB 571 LQLTWVGIKQLOARVLAVERYLKDQOLGIGWCSGKLICTTAVPWNASWNSKSLDRITWN 630
QY 524 MTWMEEREIDNYSEIYTLIEESONQOEKNEBELLELDKASIMNFDITNMLY 579
DB 631 MTWMEEREIDNYSEIYTLIEESONQOEKNEBELLELDKASIMNFDITNMLY 686

RESULT 15

US-08-472-240A-7

Sequence 7, Application US/08472240A

Patent No. 6284248

GENERAL INFORMATION:

APPLICANT: KIENY, Marie-Paule

TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/472,240A

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/956,483

FILING DATE: 31-DEC-1992

ATTORNEY/AGENT INFORMATION:

NAME: Teskin, Robin L.

REGISTRATION NUMBER: 35,030

REFERENCE/DOCKET NUMBER: 017753-055

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-6620

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:
LENGTH: 861 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Protein
LOCATION: 1..831
US-08-472-240A-7

US-08-472-240A-7

Query Match 87.1%; Score 2724.5; DB 3; Length 861;
Best Local Similarity 79.6%; Pred. No. 5,5e-227;
Matches 522; Conservative 25; Mismatches 30; Indels 79; Gaps 6;

QY 2 EKLMTVYGVGPVWKAEATTLTFCASDAKAVDTFVHNWATHACVPTDPNPQEVLENVTE 61
DB 32 EKLMTVYGVGPVWKAEATTLTFCASDAKAVDTFVHNWATHACVPTDPNPQEVLENVTE 91
62 HFMNMKNNVQEQWEDIIISLMDQSLKPCVLTPLCGA----- 98
92 NFMNMKNNDVVEQWEDIIISLMDQSLKPCVLTPLCGA----- 151
QY 99 -----GCDISVITQACP 110
DB 152 MMEKEIKKCSFNISTSIKGVQKEVAFYKLDIIPINDTTSYTLTSCNTSVITQACP 211
QY 111 KISFEPPIHYCAPAGFATLKCNKDTFNGKGPCKNVSTVOCTGIRPVVSTOLLNGSLA 170
DB 212 KVSFEPPIHYCAPAGFATLKCNKDTFNGKGPCKNVSTVOCTGIRPVVSTOLLNGSLA 271
QY 171 EEEVVRSDNFNTNAKTIIVQLKESVEINCTRPNNTRKSIH--GPGRAFYTTEIGD 228
DB 272 EEEVVRSDNFNTNAKTIIVQLKESVEINCTRPNNTRKSIH--GPGRAFYTTEIGD 330
QY 229 IROAHGNIIRAKNNTLTKQIVIKLRQF--ENKTIIVNHSGGDEIVMHSFNGGSEFFYC 287
DB 331 IROAHGNIIRAKNNTLTKQIVIKLRQF--ENKTIIVNHSGGDEIVMHSFNGGSEFFYC 390
QY 288 NSTQLFNSTWNN--NTEGSSNTEG--NTTLPRIKQIIMNMOEVGKAMYPPIRGQIRG 343
DB 391 NSTQLFNSTWNN--NTEGSSNTEG--NTTLPRIKQIIMNMOEVGKAMYPPIRGQIRG 450
QY 344 SSNTGLLTLTRDGINENGTEIFRPGGDMRDNRSELYKVKVLEPLGVAPTCKRRY 403
DB 451 SSNTGLLTLTRDGINENGTEIFRPGGDMRDNRSELYKVKVLEPLGVAPTCKRRY 510
QY 404 VOREKRAVGIGAVFLGFLGAAGSTMGASWTLTVQARLLSGIVQOQNNLRAIEAQRM 463
DB 511 VOREKRAVGIGAVFLGFLGAAGSTMGASWTLTVQARLLSGIVQOQNNLRAIEAQRM 570
QY 464 LQUTWGIKQLQARVAVERYLADQDLGIGCGSGKLICTAVPMAWSNKSLEDRINN 523
DB 571 LQUTWGIKQLQARVAVERYLADQDLGIGCGSGKLICTAVPMAWSNKSLEDRINN 630
QY 524 MTMMEWEREIDNYTSEIYTLIEESQNOQEKNEOELBLDKMASLMMFPDITWLMY 579
DB 631 MTMMEWEREIDNYTSEIYTLIEESQNOQEKNEOELBLDKMASLMMFPDITWLMY 686

Search completed: December 12, 2003, 12:34:28
Job time : 15.1662 secs

THIS PAGE BLANK (US.110)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 12:32:44 ; Search time 24.1381 Seconds
(without alignments)
4461.192 Million cell updates/sec

Title: US-10-032-162-15

Perfect score: 3129
Sequence: 1 VEKLMVTYYGVVPWKKEATT.....ELDKWASIMWFITNMLMY 579

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

1 number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3129	100.0	579	US-10-032-162-15	Sequence 15, Appl
2	3056	97.7	643	US-10-032-162-13	Sequence 17, Appl
3	2923	93.4	625	US-10-032-162-17	Sequence 1, Appl
4	2771	88.6	619	US-09-891-609-4	Sequence 4, Appl
5	2771	88.6	646	US-09-891-609-2	Sequence 2, Appl
6	2764	88.3	842	US-10-190-435-2	Sequence 2, Appl
7	2764	88.3	842	US-10-241-009-2	Sequence 2, Appl
8	2764	88.3	842	US-10-190-434B-2	Sequence 2, Appl
9	2764	88.3	842	US-10-190-305A-2	Sequence 2, Appl
10	2764	88.3	847	US-09-476-242-2	Sequence 2, Appl
11	2724.5	87.1	861	US-10-026-741-103	Sequence 103, App
12	2702	86.4	856	US-09-476-242-1	Sequence 1, Appl
13	2694	86.1	856	US-10-196-515-11	Sequence 11, Appl
14	2672.5	85.4	726	US-10-196-515-3	Sequence 3, Appl
15	2651.5	84.7	759	US-10-196-515-12	Sequence 12, Appl

16	2641.5	84.4	853	12	US-10-286-332A-33	Sequence 33, Appl
17	2641.5	84.4	853	12	US-10-280-915-33	Sequence 33, Appl
18	2641.5	84.4	853	14	US-10-003-035-33	Sequence 33, Appl
19	2625.5	83.9	868	10	US-09-938-406-1	Sequence 1, Appl
20	2535	81.0	1101	12	US-10-286-332A-53	Sequence 53, Appl
21	2535	81.0	1101	12	US-10-280-915-53	Sequence 53, Appl
22	2535	81.0	1101	14	US-10-003-035-53	Sequence 53, Appl
23	2520	80.5	1186	12	US-10-286-332A-55	Sequence 55, Appl
24	2520	80.5	1186	12	US-10-280-915-55	Sequence 55, Appl
25	2520	80.5	1186	14	US-10-003-035-55	Sequence 55, Appl
26	2495.5	79.8	860	12	US-10-190-435-6	Sequence 6, Appl
27	2495.5	79.8	860	12	US-10-241-009-6	Sequence 6, Appl
28	2495.5	79.8	860	12	US-10-190-434B-6	Sequence 6, Appl
29	2495.5	79.8	860	12	US-10-190-305A-6	Sequence 6, Appl
30	2486	79.5	858	12	US-10-190-435-150	Sequence 150, App
31	2457	78.5	867	12	US-10-190-435-3	Sequence 3, Appl
32	2457	78.5	867	12	US-10-190-435-126	Sequence 126, App
33	2457	78.5	867	12	US-10-241-009-3	Sequence 3, Appl
34	2457	78.5	867	12	US-10-190-434B-3	Sequence 3, Appl
35	2457	78.5	867	12	US-10-190-305A-3	Sequence 3, Appl
36	2449	78.3	870	12	US-10-190-435-147	Sequence 147, App
37	2448.5	78.3	803	12	US-10-190-435-135	Sequence 135, App
38	2447.5	78.2	803	12	US-10-190-435-134	Sequence 134, App
39	2447	78.2	861	12	US-10-190-435-139	Sequence 139, App
40	2446	78.2	844	10	US-09-991-258-19	Sequence 19, Appl
41	2442.5	78.1	845	12	US-10-190-435-129	Sequence 129, App
42	2442.5	78.1	845	12	US-10-190-435-130	Sequence 130, App
43	2440.5	78.0	845	12	US-10-190-435-143	Sequence 143, App
44	2439	77.9	869	12	US-10-190-435-4	Sequence 4, Appl
45	2439	77.9	869	12	US-10-241-009-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-10-032-162-15
; Sequence 15, Application US/10032162
; Publication No US20030052839A1
GENERAL INFORMATION:
; APPLICANT: BINLEY, JAMES M
; APPLICANT: SCHUEBLE, NORBERT
; APPLICANT: OLSON, WILLIAM C
; APPLICANT: PAUL, MADON J
; APPLICANT: JOHN, MOORE P
TITLE OF INVENTION: STABILIZED VIRAL ENVELOPE PROTEINS AND USES THEREOF
FILE REFERENCE: 2048/5931a2
CURRENT APPLICATION NUMBER: US/10/032,162
CURRENT FILING DATE: 2002-10-24
PRIOR APPLICATION NUMBER: 09/602,864
PRIOR FILING DATE: 2000-06-23
NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-10-032-162-15
Query Match 100.0%; Score 3129; DB 15; Length 579;
Best Local Similarity 100.0%; Pred. No. 6.9e-287;
Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VEKLMVTYYGVVPWKKEATTLTFCASDAKAYDTEVHNWATHACVPTDPNQEVLLENT 60
DB 1 VEKLMVTYYGVVPWKKEATTLTFCASDAKAYDTEVHNWATHACVPTDPNQEVLLENT 60
QY EHFMKKNWKNWQKQEDTSLMDOSLKPVCVLTPLCGAGCOTSVITACPKISPEPIYH 120
DB EHFMKKNWKNWQKQEDTSLMDOSLKPVCVLTPLCGAGCOTSVITACPKISPEPIYH 120
QY 121 YCAGAGFALCNKRTFNGKPGCKRNVSTVQCTHGIRPVSSTOLLNGSLAEVEVIRSDN 180

Db 121 YCAPAGFAILKCNKDTFNKGKPCKNVSTVQCTHGIRPVVSTOLLNGLSLAEVEVIRSDN 180
Qy 181 FTNNAKTIIVOLKESVEINCTRPNNNTKRSIHIGPRAFYTTGEIIGDIRAHONISRAK 240
Db 181 FTNNAKTIIVOLKESVEINCTRPNNNTKRSIHIGPRAFYTTGEIIGDIRAHONISRAK 240
Qy 241 WNDLTKQIVIKLREQFENKTIIVFNHSSGGDEPEIIVHNSFCGEFFYCNSITOLFNSITNNN 300
Db 241 WNDLTKQIVIKLREQFENKTIIVFNHSSGGDEPEIIVHNSFCGEFFYCNSITOLFNSITNNN 300
Qy 301 TEGSNNTEGNTITLPCRKOIINMGOEVKAMVAPPIRGQIRCSSNITGLLTRDGGINE 360
Db 301 TEGSNNTEGNTITLPCRKOIINMGOEVKAMVAPPIRGQIRCSSNITGLLTRDGGINE 360
Qy 361 NGTEIFRPGGDMRSELYKVKVTEIPLGVAATPKCKRRVQREKRAVIGAVLGF 420
Db 361 NGTEIFRPGGDMRSELYKVKVTEIPLGVAATPKCKRRVQREKRAVIGAVLGF 420
Qy 421 LGAAGSTMGASMTLTVQARLLLSGIYQOONNLRAIEAOQRMQLTWGIKQLOARVLA 480
Db 421 LGAAGSTMGASMTLTVQARLLLSGIYQOONNLRAIEAOQRMQLTWGIKQLOARVLA 480
Qy 481 VERYLQDQOLLGIWCGSKLICCTAVPWNASWSNKSIDRIWNNMTWEMEREIDNYTSEI 540
Db 481 VERYLQDQOLLGIWCGSKLICCTAVPWNASWSNKSIDRIWNNMTWEMEREIDNYTSEI 540
Qy 541 YTLIESONOEKNEOELELDKMASIMNMPDITNLMY 579
Db 541 YTLIESONOEKNEOELELDKMASIMNMPDITNLMY 579

RESULT 2

US-10-032-162-13
; Sequence 13, Application US/10032162
; Publication No. US20030052839A1
; GENERAL INFORMATION:
; APPLICANT: BINLEY, JAMES M
; APPLICANT: SCHUEBKE, NOBERT
; APPLICANT: OLSON, WILLIAM C
; APPLICANT: PAUL, MADDON J
; APPLICANT: JOHN, MOORE P
; TITLE OF INVENTION: STABILIZED VIRAL ENVELOPE PROTEINS AND USES THEREOF
; FILE REFERENCE: 2048/59331az
; CURRENT APPLICATION NUMBER: US/10/032,162
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/602,864
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 643
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-10-032-162-13

Query Match 97.7%; Score 3056; DB 15; Length 643;
Best Local Similarity 89.3%; Pred. No. 6,5e-280;
Matches 574; Conservative 0; Mismatches 5; Indels 64; Gaps 1;
Qy 1 VEKLMVTYYGVVWPKKATTTLFCASDAKAYDTEVHNWMAHACVPTDPNPOEVLENT 60
Db 1 VEKLMVTYYGVVWPKKATTTLFCASDAKAYDTEVHNWMAHACVPTDPNPOEVLENT 60
Qy 61 EHFNNKNNKNNWQOQEDIIISLMDOSLKPCVKLTPLCVLJNCKDVNATNTNDSGTMERG 98
Db 61 EHFNNKNNKNNWQOQEDIIISLMDOSLKPCVKLTPLCVLJNCKDVNATNTNDSGTMERG 120
Qy 99 -----GCDTSVITQACPKISFEP 116
Db 121 EIKKCSFNTTISIRDEVQKVALFKLDVVPIDNNNTSYRLISCTSYITQACPKISFEP 180
Qy 117 IPIHYCAPAGFAILKCNKDTFNKGKPCKNVSTVQCTHGIRPVVSTOLLNGLSLAEVEV 176
Db 117 IPIHYCAPAGFAILKCNKDTFNKGKPCKNVSTVQCTHGIRPVVSTOLLNGLSLAEVEV 176

Db 181 IPIHYCAPAGFAILKCNKDTFNKGKPCKNVSTVQCTHGIRPVVSTOLLNGLSLAEVEV 240
Qy 177 RSDNFTNNAKTIIVOLKESVEINCTRPNNNTKRSIHIGPRAFYTTGEIIGDIRAHON 236
Db 241 RSDNFTNNAKTIIVOLKESVEINCTRPNNNTKRSIHIGPRAFYTTGEIIGDIRAHON 300
Qy 237 SRAKMDTLKQIVIKLREQFENKTIIVFNHSSGGDEPEIIVHNSFCGEFFYCNSITOLFNS 236
Db 301 SRAKMDTLKQIVIKLREQFENKTIIVFNHSSGGDEPEIIVHNSFCGEFFYCNSITOLFNS 360
Qy 297 WNNTEGNNTEGNTITLPCRKOIINMGOEVKAMVAPPIRGQIRCSSNITGLLTRD 356
Db 361 WNNTEGNNTEGNTITLPCRKOIINMGOEVKAMVAPPIRGQIRCSSNITGLLTRD 420
Qy 357 GINENGTEIFRPGGDMRSELYKVKVTEIPLGVAATPKCKRRVQREKRAVIGAV 416
Db 421 GINENGTEIFRPGGDMRSELYKVKVTEIPLGVAATPKCKRRVQREKRAVIGAV 480
Qy 417 FLGFLGAAGSTMGASMTLTVQARLLLSGIYQOONNLRAIEAOQRMQLTWGIKQLOA 476
Db 481 FLGFLGAAGSTMGASMTLTVQARLLLSGIYQOONNLRAIEAOQRMQLTWGIKQLOA 540
Qy 477 RVIAVERYLDQOLLGIWCGSKLICCTAVPWNASWSNKSIDRIWNNMTWEMEREIDNY 536
Db 541 RVIAVERYLDQOLLGIWCGSKLICCTAVPWNASWSNKSIDRIWNNMTWEMEREIDNY 600
Qy 537 TSELYTLIESONOEKNEOELELDKMASIMNMPDITNLMY 579
Db 601 TSELYTLIESONOEKNEOELELDKMASIMNMPDITNLMY 643

RESULT 3

US-10-032-162-17
; Sequence 17, Application US/10032162
; Publication No. US20030052839A1
; GENERAL INFORMATION:
; APPLICANT: BINLEY, JAMES M
; APPLICANT: SCHUEBKE, NOBERT
; APPLICANT: OLSON, WILLIAM C
; APPLICANT: PAUL, MADDON J
; APPLICANT: JOHN, MOORE P
; TITLE OF INVENTION: STABILIZED VIRAL ENVELOPE PROTEINS AND USES THEREOF
; FILE REFERENCE: 2048/59331az
; CURRENT APPLICATION NUMBER: US/10/032,162
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/602,864
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 625
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (151)..(151)
; OTHER INFORMATION: X=UNKNOWN AMINO ACID
; NAME/KEY: MISC FEATURE
; LOCATION: (205)..(205)
; OTHER INFORMATION: X=UNKNOWN AMINO ACID
; NAME/KEY: MISC FEATURE
; LOCATION: (213)..(213)
; OTHER INFORMATION: X=UNKNOWN AMINO ACID
US-10-032-162-17

Query Match 93.4%; Score 2923; DB 15; Length 625;
Best Local Similarity 86.0%; Pred. No. 2.3e-267;
Matches 553; Conservative 0; Mismatches 8; Indels 82; Gaps 2;
Qy 1 VEKLMVTYYGVVWPKKATTTLFCASDAKAYDTEVHNWMAHACVPTDPNPOEVLENT 60
Db 1 VEKLMVTYYGVVWPKKATTTLFCASDAKAYDTEVHNWMAHACVPTDPNPOEVLENT 60

Db 1 VEKLMVTYYGVVPWKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENVT 60
 Qy 61 EHFNNMKNMNVQOMEDIIISLMDOSLKPVCVLTPLCGA----- 98
 Db 61 EHFNNMKNMNVQOMEDIIISLMDOSLKPVCVLTPLCVTLNCKDVANTNTNDEGMENG 120
 Qy 99 -----GCDTSVITQACPISFEP 116
 Db 121 EIKNCSFNITTSIRDEVQKEVALFYKLDVYXIDNNNTSYRLISCDTSVITQACPISFEP 180
 Qy 117 IPIHCAAPGAFALICNDKTFNGKPCCKNVSTVQCTHGRPVYSTOLLNGSLAEBEVY 176
 Db 181 IPIHCAAPGAFALICNDKTFNGKPCCKNVSTVQCTHGRPVYSTOLLNGSLAEBEVY 240
 Qy 177 RSDNFTNNAKTIIVOLKESVEINCTRPNNTRKSHIGGRAFYTTGELIGIROAHCHI 236
 Db 241 RSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGIRQAHCHI 282
 Qy 237 SRAKNDTLKOIIVIKLREQFENKTIIVFNHSGGDEBEIVMHSFNCGGEFFYCNSDTLPNST 296
 Db 283 SRAKNDTLKOIIVIKLREQFENKTIIVFNHSGGDEBEIVMHSFNCGGEFFYCNSDTLPNST 342
 Qy 297 WNNNTEGNNTEGNTITTLPCRILKOIINMVEGKAMYPPIRGQIRCSNITGLLTRDG 356
 Db 343 WNNNTEGNNTEGNTITTLPCRILKOIINMVEGKAMYPPIRGQIRCSNITGLLTRDG 402
 Qy 357 GINENGTEIFRPGGDMRDMRSELYKYKVYKIEPIGVAFTCKRVRVOREKRAVGIGAV 416
 Db 403 GINENGTEIFRPGGDMRDMRSELYKYKVYKIEPIGVAFTCKRVRVOREKRAVGIGAV 462
 Qy 417 FLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOONNLLRAIEAQORMLQITVWGIKOLOA 476
 Db 463 FLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOONNLLRAIEAQORMLQITVWGIKOLOA 522
 Qy 477 RVLAVERYLGDQOOLLGIGWSSGKLICTTAVPNMASSNSKSLDIRINNNTMMEWEIEDNY 536
 Db 523 RVLAVERYLGDQOOLLGIGWSSGKLICTTAVPNMASSNSKSLDIRINNNTMMEWEIEDNY 582
 Qy 537 TSEIYTLIEESONQOEKNEOELLEBDKMASLMMWFDTNNMLMY 579
 Db 583 TSEIYTLIEESONQOEKNEOELLEBDKMASLMMWFDTNNMLMY 625

RESULT 4
 US-09-891-609-4
 ; Sequence 4, Application US/09891609
 ; Patent No. US20020127238A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Stamatatos, Leonidas
 ; TITLE OF INVENTION: HIV-1 VACCINES AND SCREENING METHODS THEREFOR
 ; FILE REFERENCE: 2570-1-002N
 ; CURRENT APPLICATION NUMBER: US/09/891,609
 ; CURRENT FILING DATE: 2001-06-26
 ; PRIOR APPLICATION NUMBER: 60/214,608
 ; PRIOR FILING DATE: 2000-06-27
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 4
 ; LENGTH: 619
 ; TYPE: PRT
 ; ORGANISM: Human immunodeficiency virus type 1
 US-09-891-609-4

Query Match 88.6%; Score 2771; DB 10; Length 619;
 Best Local Similarity 85.6%; Pred. No. 5,5e-253;
 Matches 529; Conservative 16; Mismatches 31; Indels 42; Gaps 6;

Qy 1 VEKLMVTYYGVVPWKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENVT 60
 Db 3 VEKLMVTYYGVVPWKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENVT 62
 Qy 61 EHFNNMKNMNVQOMEDIIISLMDOSLKPVCVLTPLC----- 96
 Db 1 EHFNNMKNMNVQOMEDIIISLMDOSLKPVCVLTPLC----- 96

Db 63 EHFNNMKNMNVQOMEDIIISLMDOSLKPVCVLTPLCVTLHCTNLKNAATNTKSSNNKEMDR 122
 Qy 97 -----GAG-----CDTSVITQACPISFEPPIHYCAPAGFALLKNDKTFNGKCP 142
 Db 123 GEIKNCSFKVAGAKINCNTSVITQACPVSEFPIPIHYCAPAGFALLKNDKTFNGKCP 182
 Qy 143 CKNVSTVQCTHGRPVYSTOLLNGSLAEBEVIRSDNFTNNAKTIIVOLKESVEINCTR 202
 Db 183 CKNVSTVQCTHGRPVYSTOLLNGSLAEBEVIRSENFPTNNAKTIIVOLKESVEINCTR 242
 Qy 203 PNNNTRKSHIGGRAFYTTGELIGIROAHCHNISRAKNDTLKOIIVIKLREQFENKTIIV 262
 Db 243 PNNNTRKSHIGGRAFYTTGELIGIROAHCHNISRAKNDTLKOIIVIKLREQFENKTIIV 302
 Qy 263 FHNHSGGDEBEIVMHSFNCGGEFFYCNSDTLPNSTNNNTGEGNTITTLPCRILKOI 322
 Db 303 FHNHSGGDEBEIVMHSFNCGGEFFYCNSDTLPNSTNNNTGEGNTITTLPCRILKOI 360
 Qy 323 NMMQEVGKAMYPPIRGQIRCSNITGLLTRDGGIN-ENGTEIFRPGGDMRDMRSEL 381
 Db 361 NMMQEVGKAMYPPIRGQIRCSNITGLLTRDGKEISNTTEIFRPGGDMRDMRSEL 420
 Qy 382 YKYKVYKIEPIGVAFTCKRVRVOREKRAVGIGAVFLGFLGAAGSTWGAASMTLTVQARL 441
 Db 421 YKYKVYKIEPIGVAFTCKRVRVOREKRAVGIGAVFLGFLGAAGSTWGAASMTLTVQARL 480
 Qy 442 LLSGIVQOONNLLRAIEAQORMLQITVWGIKOLOARVLAVERYLGDQOOLLGIGWSSGKLI 501
 Db 481 LLSGIVQOONNLLRAIEAQORMLQITVWGIK-LQARVLAVERYLKDQOOLLGIGWSSGKLI 539
 Qy 502 CTTAVPNMASSNSKSLDIRINNNTMMEWEIEDNTSYSEIYTLIEESONQOEKNEOELLE 561
 Db 540 CTTAVPNMASSNSKSLDIRINNNTMMEWEIEDNTSYSEIYTLIEESONQOEKNEOELLE 599
 Qy 562 DKMASLMMWFDTNNMLMY 579
 Db 600 DKMASLMMWFDTNNMLMY 617

RESULT 5
 US-09-891-609-2
 ; Sequence 2, Application US/09891609
 ; Patent No. US20020127238A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Stamatatos, Leonidas
 ; TITLE OF INVENTION: HIV-1 VACCINES AND SCREENING METHODS THEREFOR
 ; FILE REFERENCE: 2570-1-002N
 ; CURRENT APPLICATION NUMBER: US/09/891,609
 ; CURRENT FILING DATE: 2001-06-26
 ; PRIOR APPLICATION NUMBER: 60/214,608
 ; PRIOR FILING DATE: 2000-06-27
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 2
 ; LENGTH: 646
 ; TYPE: PRT
 ; ORGANISM: Human immunodeficiency virus type 1
 US-09-891-609-2

Query Match 88.6%; Score 2771; DB 10; Length 646;
 Best Local Similarity 85.6%; Pred. No. 5,8e-253;
 Matches 529; Conservative 16; Mismatches 31; Indels 42; Gaps 6;

Qy 1 VEKLMVTYYGVVPWKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENVT 60
 Db 30 VEKLMVTYYGVVPWKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENVT 89
 Qy 61 EHFNNMKNMNVQOMEDIIISLMDOSLKPVCVLTPLC----- 96
 Db 90 EHFNNMKNMNVQOMEDIIISLMDOSLKPVCVLTPLCTTLHCTNLKNAATNTKSSNNKEMDR 149
 Qy 97 -----GAG-----CDTSVITQACPISFEPPIHYCAPAGFALLKNDKTFNGKCP 142
 Db 1 EHFNNMKNMNVQOMEDIIISLMDOSLKPVCVLTPLC----- 96

```

Db 150 GEIKNCSEFKVAGAGKLINCNTSVITQACPVSFEPIPIHYCAPAGFALIKCNCKKFNKSGCP 209
Qy 143 CKNVSTVOCTHGIRPVYSTOLLNGSLAESEVIRSDNFTNNAKTIIVOLKESEVINCTR 202
Db 210 CTNVSTVOCTHGIRPVYSTOLLNGSLAESEVIRSENFTDNAKTIIVOLKESEVINCTR 269
Qy 203 PNNTRKSIHIGPRAFYTTGEIIGDIRQAHCHNISRANKNDLKOIYKLEPOENKTIIV 262
Db 270 PNNTRKSIITGPRAFYATGDIIGDIRQAHCHNISGEKMNNTLKOIYKLEPOENKTIIV 329
Qy 263 FNHSSGDEPEIVHMSFNGCGEFPYCNSTOLFNSTNNNTBESNTEGNTITLPCRKOI 322
Db 330 FKOSGSDPEIVHMSFNGCGEFPYCNSTOLFNSTNNTTGPNNNTG-TITLPCRKOI 387
Qy 323 NNMGEVGAAPYPIRQIRCSSNITGLLTRDGIN-ENGTEIFRPGGDMRDNRSSEL 381
Db 388 NRMGEVGAAPYPIRQIRCSSNITGLLTRDGEKISNTTEIFRPGGDMRDNRSSEL 447
Qy 382 YKYKVKIPELVAPTKCKRRVOREKRAVGIGAVFLGFLGAAGSTWGAASMTLTVQARL 441
Db 448 YKYKVKIPELVAPTKCKRRVOREKRAVTLGAMFLGFLGAAGSTWGAASMTLTVQARQ 507
Qy 442 LLSGIVQOQNNLLRAIEAQOQMLQTVGIRKQOLAVRYIGDQQLGIMGSGKLI 501
Db 508 LLSGIVQOQNNLLRAIEAQOHLQTLTWGIR-LQARVLAVERYLKQQLGIMGSGKLI 566
Qy 502 CCTAVPNNASWSKSLDRIANNMTMWEEREIDNTSYTLIEESONQOEKNEOELTEL 561
Db 567 CTIAVPNNASWSKSLDRIANNMTMWEEREIDNTSYTLIEESONQOEKNEOELTEL 626
Qy 562 DKWASLNMWPDITNMLWY 579
Db 627 DKWASLNMWPDITNMLWY 644

```

RESULT 6 US-10-190-435-2

```

; Sequence 2, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; FILE REFERENCE: P18133.003 / 2302-18133
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 842
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SFI62
US-10-190-435-2

```

```

Query Match 88.3%; Score 2764; DB 12; Length 842;
Best Local Similarity 81.7%; Pred. No. 4e-252;
Matches 527; Conservative 16; Mismatches 34; Indels 68; Gaps 4;
Qy 1 VEKLMVTYVYGVVPWKKAATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENTV 60
Db 25 VEKLMVTYVYGVVPWKKAATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENTV 84
Qy 61 EHEFNWKNWVQOQEDIIISLMDOSLKPVCVLTPLCGA----- 98
Db 85 EHEFNWKNWVQOQEDIIISLMDOSLKPVCVLTPLCGA----- 144
Qy 99 -----GCDTSVITQACPKISFE 115

```

```

Db 145 GEIKNCSEFKVTSIRNKQKEVALFYELDVVPIDNDNTSYKLIINCNTSVITQACPVSFE 204
Qy 116 PIPHYCAPAGFALIKCNCKKFNKSGCPKRVSTVOCTHGIRPVYSTOLLNGSLAESEV 175
Db 205 PIPHYCAPAGFALIKCNCKKFNKSGCPKRVSTVOCTHGIRPVYSTOLLNGSLAESEV 264
Qy 176 IRSNFTNNAKTIIVOLKESEVINCTRPNNNTKSIHIGPRAFYTTGEIIGDIRQAHCHN 235
Db 265 IRSNFTNNAKTIIVOLKESEVINCTRPNNNTKSIITGPRAFYATGDIIGDIRQAHCHN 324
Qy 236 ISRAKNDTLKOIYKLEPOENKTIIVFNHSSGDEPEIVHMSFNGCGEFPYCNSTOLFNS 295
Db 325 ISGEKMNNTLKOIYKLEPOENKTIIVFNHSSGDEPEIVHMSFNGCGEFPYCNSTOLFNS 384
Qy 296 TNNNTBESNTEGNTITLPCRKOIINNMGEVGAAPYPIRQIRCSSNITGLLTRD 355
Db 385 TW-NNTTGPNNNTG-TITLPCRKOIINRMGEVGAAPYPIRQIRCSSNITGLLTRD 442
Qy 356 GGIN-ENGTEIFRPGGDMRDNRSSELYKYKVKIPELVAPTKCKRRVOREKRAVGIG 414
Db 443 GGEKISNTTEIFRPGGDMRDNRSSELYKYKVKIPELVAPTKCKRRVOREKRAVTLG 502
Qy 415 AVFLGFLGAAGSTWGAASMTLTVQARLLSGIVQOQNNLLRAIEAQOQMLQTVGIRKOL 474
Db 503 AMFLGFLGAAGSTWGAASMTLTVQARQLLSGIVQOQNNLLRAIEAQOHLQTLTWGIRKOL 562
Qy 475 QARVLAVERYLKQQLGIMGSGKLICTTAVPNNASWSKSLDRIANNMTMWEEREID 534
Db 563 QARVLAVERYLKQQLGIMGSGKLICTTAVPNNASWSKSLDRIANNMTMWEEREID 622
Qy 535 NYTSEIYTLIEESONQOEKNEOELTELDKWASLNMWPDITNMLWY 579
Db 623 NYTSEIYTLIEESONQOEKNEOELTELDKWASLNMWPDITNMLWY 667

```

RESULT 7 US-10-241-009-2

```

; Sequence 2, Application US/10241009
; Publication No. US20030170614A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B
; FILE REFERENCE: 2300-1621.21
; CURRENT FILING DATE: 2002-12-13
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 842
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SFI62
US-10-241-009-2

```

```

Query Match 88.3%; Score 2764; DB 12; Length 842;
Best Local Similarity 81.7%; Pred. No. 4e-252;
Matches 527; Conservative 16; Mismatches 34; Indels 68; Gaps 4;
Qy 1 VEKLMVTYVYGVVPWKKAATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENTV 60
Db 25 VEKLMVTYVYGVVPWKKAATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENTV 84
Qy 61 EHEFNWKNWVQOQEDIIISLMDOSLKPVCVLTPLCGA----- 98
Db 85 EHEFNWKNWVQOQEDIIISLMDOSLKPVCVLTPLCGA----- 144
Qy 99 -----GCDTSVITQACPKISFE 115

```



```

Db      145 GEIKNSCFKVTTSIRNKQKEVALFYKLDVVPIDNDNTSYKLINCNTSVITQACPKVSFE 204
Qy      116 PIPHYCAPAGFAILKNDKTEFGNGKPCKNVSTVOCTHGIRPVVSTOLLNGLAEVY 175
Db      205 PIPHYCAPAGFAILKNDKTEFGNGKPCKNVSTVOCTHGIRPVVSTOLLNGLAEVY 264
Qy      176 IRSDFPNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPRAFYTTGEEIGDIRQAHCN 235
Db      265 IRSEFPTNDAKTIIVQLKESVEINCTRPNNNTRKSIHIGPRAFYATGDIIDIRQAHCN 324
Qy      226 ISRAKNDTLKOIIVKLAEQFENKTIIVPHSSGGDEPEIYVMSFNGGEPFYCNSTOLFNS 295
Db      325 ISGEKMNNTLKOIIVKLAEQFENKTIIVPHSSGGDEPEIYVMSFNGGEPFYCNSTOLFNS 384
Qy      226 TNNNTGSSNNTGNTITLPCRIKOIIMMOEVRKAMYAPPIRGQIRCSNITGLLTFRD 355
Db      385 TW-ANTIGPNNTNG-TITLPCRIKOIIMMOEVRKAMYAPPIRGQIRCSNITGLLTFRD 442
Qy      356 GGIN-ENGTEIFRPGGDMRDNRSELYKYVVKIEPLGVAFTKCRVVOREKRAVGIG 414
Db      443 GKEISNTTEIFRPGGDMRDNRSELYKYVVKIEPLGVAFTKCRVVOREKRAVTLG 502
Qy      415 AVFLGFLGAAGSTWGAASMTLTVQARLLSGIVQOQNNLLRAIEAQOQMLQLTWVGIKQL 474
Db      503 AMFLGFLGAAGSTWGAASMTLTVQARLLSGIVQOQNNLLRAIEAQOQMLQLTWVGIKQL 562
Qy      475 QARVLAVERYLGDQDLGIMGSGKLICTAVPNNASNSKSLDRIINNMTMWEWEREID 534
Db      563 QARVLAVERYLKDQDLGIMGSGKLICTAVPNNASNSKSLDRIINNMTMWEWEREID 622
Qy      535 NYTSEIYTLIESQNOQEKNEOELELDKMASLNMWFDITNMLWY 579
Db      623 NYTNLIYTLIESQNOQEKNEOELELDKMASLNMWFDISKMLWY 667

```

```

RESULT 8
US-10-190-434B-2
; Sequence 2, Application US/10190434B
; Publication No. US20030194800A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARRETT, Susan
; APPLICANT: LIAN, Ying
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B
; FILE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 2300-1621.20
; CURRENT APPLICATION NUMBER: US/10/190,434B
; CURRENT FILING DATE: 2002-07-05
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 842
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SP162
US-10-190-434B-2

```

```

Query Match      88.3%; Score 2764; DB 12; Length 842;
Best Local Similarity 81.7%; Pred. No. 4e-252;
Matches 527; Conservative 16; Mismatches 34; Indels 68; Gaps 4;
Qy      1 VEKLVVTVYGVVWKEATTTLPCASDAKAYDTEVHNVWATACVPTDPNPOEVYLENT 60
Db      25 VEKLVVTVYGVVWKEATTTLPCASDAKAYDTEVHNVWATACVPTDPNPOEVYLENT 84
Qy      61 EHFNNMKNMNVQOMQEDIIISLMDQSLKPCVKLTPLCGA----- 98
Db      85 EHFNNMKNMNVQOMQEDIIISLMDQSLKPCVKLTPLCVTLHCTNLKNAITNTSSNNKEMDR 144
Qy      99 -----GCDTSVITQACPKVSFE 115
Db      145 GEIKNSCFKVTTSIRNKQKEVALFYKLDVVPIDNDNTSYKLINCNTSVITQACPKVSFE 204

```

```

Qy      116 PIPHYCAPAGFAILKNDKTEFGNGKPCKNVSTVOCTHGIRPVVSTOLLNGLAEVY 175
Db      205 PIPHYCAPAGFAILKNDKTEFGNGKPCKNVSTVOCTHGIRPVVSTOLLNGLAEVY 264
Qy      176 IRSDFPNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPRAFYTTGEEIGDIRQAHCN 235
Db      265 IRSEFPTNDAKTIIVQLKESVEINCTRPNNNTRKSIHIGPRAFYATGDIIDIRQAHCN 324
Qy      226 ISRAKNDTLKOIIVKLAEQFENKTIIVPHSSGGDEPEIYVMSFNGGEPFYCNSTOLFNS 295
Db      325 ISGEKMNNTLKOIIVKLAEQFENKTIIVPHSSGGDEPEIYVMSFNGGEPFYCNSTOLFNS 384
Qy      226 TNNNTGSSNNTGNTITLPCRIKOIIMMOEVRKAMYAPPIRGQIRCSNITGLLTFRD 355
Db      385 TW-ANTIGPNNTNG-TITLPCRIKOIIMMOEVRKAMYAPPIRGQIRCSNITGLLTFRD 442
Qy      356 GGIN-ENGTEIFRPGGDMRDNRSELYKYVVKIEPLGVAFTKCRVVOREKRAVGIG 414
Db      443 GKEISNTTEIFRPGGDMRDNRSELYKYVVKIEPLGVAFTKCRVVOREKRAVTLG 502
Qy      415 AVFLGFLGAAGSTWGAASMTLTVQARLLSGIVQOQNNLLRAIEAQOQMLQLTWVGIKQL 474
Db      503 AMFLGFLGAAGSTWGAASMTLTVQARLLSGIVQOQNNLLRAIEAQOQMLQLTWVGIKQL 562
Qy      475 QARVLAVERYLGDQDLGIMGSGKLICTAVPNNASNSKSLDRIINNMTMWEWEREID 534
Db      563 QARVLAVERYLKDQDLGIMGSGKLICTAVPNNASNSKSLDRIINNMTMWEWEREID 622
Qy      535 NYTSEIYTLIESQNOQEKNEOELELDKMASLNMWFDITNMLWY 579
Db      623 NYTNLIYTLIESQNOQEKNEOELELDKMASLNMWFDISKMLWY 667

```

```

RESULT 9
US-10-190-305A-2
; Sequence 2, Application US/10190305A
; Publication No. US20030198621A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARRETT, Susan
; APPLICANT: LIAN, Ying
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR
; FILE OF INVENTION: TYPE C POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 2302-18702 / 18702.002
; CURRENT APPLICATION NUMBER: US/10/190,305A
; CURRENT FILING DATE: 2002-07-05
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 842
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SP162
US-10-190-305A-2

```

```

Query Match      88.3%; Score 2764; DB 12; Length 842;
Best Local Similarity 81.7%; Pred. No. 4e-252;
Matches 527; Conservative 16; Mismatches 34; Indels 68; Gaps 4;
Qy      1 VEKLVVTVYGVVWKEATTTLPCASDAKAYDTEVHNVWATACVPTDPNPOEVYLENT 60
Db      25 VEKLVVTVYGVVWKEATTTLPCASDAKAYDTEVHNVWATACVPTDPNPOEVYLENT 84
Qy      61 EHFNNMKNMNVQOMQEDIIISLMDQSLKPCVKLTPLCGA----- 98
Db      85 EHFNNMKNMNVQOMQEDIIISLMDQSLKPCVKLTPLCVTLHCTNLKNAITNTSSNNKEMDR 144
Qy      99 -----GCDTSVITQACPKVSFE 115
Db      145 GEIKNSCFKVTTSIRNKQKEVALFYKLDVVPIDNDNTSYKLINCNTSVITQACPKVSFE 204

```

QY 116 PIPHYCAPAGFALIKCNDKTFNGKPCKNVSTVOCTHGIRPVYSTOLLNGSLAESEV 175
 DB 205 PIPHYCAPAGFALIKCNDKTFNGKPCKNVSTVOCTHGIRPVYSTOLLNGSLAESEV 264
 QY 176 IRSNFTNNAKTIIVOLKESVEINCTRNPNNTKRSIHIGPRAFYATGDIIGDIRAHCN 235
 DB 265 IRSNFTNNAKTIIVOLKESVEINCTRNPNNTKRSIHIGPRAFYATGDIIGDIRAHCN 324
 QY 236 ISRAKMDTLKQIYIKLREOPENKTIYFNHSSGGDEIIVHSPFCGGEFFYCNSDTOLPNS 295
 DB 325 ISRAKMDTLKQIYIKLREOPENKTIYFNHSSGGDEIIVHSPFCGGEFFYCNSDTOLPNS 384
 QY 296 TANNTEGSNNTEGNTITLPCRIOIIMMOEYKAMYPPIRQIICSSNITGLLTRD 355
 DB 385 TM-NTTIGPNNTNG-TITLPCRIOIIMMOEYKAMYPPIRQIICSSNITGLLTRD 442
 QY 356 GGIN-ENGTEIFRGGGDMNDRSELYKVVKIEPLGVAFTKRRVQREKRAVTLG 414
 DB 443 GGEKISNTTEIFRGGGDMNDRSELYKVVKIEPLGVAFTKRRVQREKRAVTLG 502
 QY 415 AVFLGFIAGAAGSTMGASMTLTVQARLLSGIVQOONNLLRAIEAOORMLQITWGIKOL 474
 DB 503 AVFLGFIAGAAGSTMGASMTLTVQARLLSGIVQOONNLLRAIEAOORMLQITWGIKOL 562
 QY 475 QARVLAVERYLGDQOLGIGWCSGKLICTAVPNNASMSNKSIDRIMNNTWMEEREID 534
 DB 563 QARVLAVERYLGDQOLGIGWCSGKLICTAVPNNASMSNKSIDRIMNNTWMEEREID 622
 QY 535 NYTSEIYTLIEESONQOEKNEQELLEDKXASIMNWEDITNWLWY 579
 DB 623 NYTSEIYTLIEESONQOEKNEQELLEDKXASIMNWEDITNWLWY 667

RESULT 10

US-09-476-242-2
 ; Sequence 2, Application US/09476242
 ; Patent No. US20020146683A1
 ; GENERAL INFORMATION:
 ; APPLICANT: BARNETT, Susan
 ; APPLICANT: HARTOG, Karin
 ; APPLICANT: MARTIN, Eric
 ; TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES
 ; FILE REFERENCE: 1605.002
 ; CURRENT APPLICATION NUMBER: US/09/476.242
 ; CURRENT FILING DATE: 1999-12-30
 ; NUMBER OF SEQ ID NOS: 26
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 847
 ; TYPE: PRT
 ; ORGANISM: Human immunodeficiency virus
 US-09-476-242-2

Query Match 88.3%; Score 2764; DB 10; Length 847;
 Best Local Similarity 81.7%; Pred. No. 4e-252;
 Matches 527; Conservative 16; Mismatches 34; Indels 68; Gaps 4;

QY 1 VEXIMVTVYGVVPWKEATTLTFCASDAKAYDEVHNNWATHACVPDPNQEVLNVT 60
 DB 30 VEXIMVTVYGVVPWKEATTLTFCASDAKAYDEVHNNWATHACVPDPNQEVLNVT 89
 QY 61 EHPNNNNNNNNWQOEDIIISLMDOSLKPVCVLTPLGCA----- 98
 DB 90 ENFMNNNNNNWQOEDIIISLMDOSLKPVCVLTPLGCA----- 149
 QY 99 -----GDDTSVITQACEKISFE 115
 DB 150 GEIKNCSFKVTSIRNKWQKEYALFYKLDVPIIDNDNTSYKLINCNTSVITQACEKISFE 209
 QY 116 PIPHYCAPAGFALIKCNDKTFNGKPCKNVSTVOCTHGIRPVYSTOLLNGSLAESEV 175
 DB 210 PIPHYCAPAGFALIKCNDKTFNGKPCKNVSTVOCTHGIRPVYSTOLLNGSLAESEV 269

QY 176 IRSNFTNNAKTIIVOLKESVEINCTRNPNNTKRSIHIGPRAFYATGDIIGDIRAHCN 235
 DB 270 IRSNFTNNAKTIIVOLKESVEINCTRNPNNTKRSIHIGPRAFYATGDIIGDIRAHCN 329
 QY 236 ISRAKMDTLKQIYIKLREOPENKTIYFNHSSGGDEIIVHSPFCGGEFFYCNSDTOLPNS 295
 DB 330 ISRAKMDTLKQIYIKLREOPENKTIYFNHSSGGDEIIVHSPFCGGEFFYCNSDTOLPNS 389
 QY 296 TANNTEGSNNTEGNTITLPCRIOIIMMOEYKAMYPPIRQIICSSNITGLLTRD 355
 DB 390 TM-NTTIGPNNTNG-TITLPCRIOIIMMOEYKAMYPPIRQIICSSNITGLLTRD 447
 QY 356 GGIN-ENGTEIFRGGGDMNDRSELYKVVKIEPLGVAFTKRRVQREKRAVTLG 414
 DB 448 GGEKISNTTEIFRGGGDMNDRSELYKVVKIEPLGVAFTKRRVQREKRAVTLG 507
 QY 415 AVFLGFIAGAAGSTMGASMTLTVQARLLSGIVQOONNLLRAIEAOORMLQITWGIKOL 474
 DB 508 AVFLGFIAGAAGSTMGASMTLTVQARLLSGIVQOONNLLRAIEAOORMLQITWGIKOL 567
 QY 475 QARVLAVERYLGDQOLGIGWCSGKLICTAVPNNASMSNKSIDRIMNNTWMEEREID 534
 DB 568 QARVLAVERYLGDQOLGIGWCSGKLICTAVPNNASMSNKSIDRIMNNTWMEEREID 627
 QY 535 NYTSEIYTLIEESONQOEKNEQELLEDKXASIMNWEDITNWLWY 579
 DB 628 NYTSEIYTLIEESONQOEKNEQELLEDKXASIMNWEDITNWLWY 672

RESULT 11

US-10-026-741-103
 ; Sequence 103, Application US/10026741
 ; Publication No. US20030049604A1
 ; GENERAL INFORMATION:

APPLICANT: CHARNEAU, PIERRE
 CLAVEL, FRANCOISE
 BORMAN, ANDREW
 OUILLENT, CAROLINE
 GUETARD, DENISE
 MONTAGNER, LUC
 DONON DE SAINT-MARTIN, JACQUELINE
 COHEN, JACQUES
 TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF HIV-1 TYPE (OR
 SUBTYPE) ANTIGENS
 NUMBER OF SEQUENCES: 103
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
 Dunner, L.L.P.
 STREET: 1300 I Street, N.W.
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20005-3315
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/026.741
 FILING DATE: 27-Dec-2001
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/817,441
 FILING DATE: 31-AUG-1998
 APPLICATION NUMBER: PCT/FR 95/01391
 FILING DATE: 20-OCT-1995
 APPLICATION NUMBER: FR 9412554
 FILING DATE: 20-OCT-1994
 APPLICATION NUMBER: FR 9502526
 FILING DATE: 03-MAR-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Meyers, Kenneth J.
 REGISTRATION NUMBER: 25,146

REFERENCE/DOCKET NUMBER: 03260.6005-00000
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-408-4000
 TELEFAX: 202-408-4400
 INFORMATION FOR SEQ ID NO: 103:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 861 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 103:
 US-10-026-741-103

Query Match 87.1%; Score 2724.5; DB 15; Length 861;
 Best Local Similarity 79.6%; Pred. No. 2,2e-248;
 Matches 522; Conservative 25; Mismatches 30; Indels 79; Gaps 6;

```

QY 2 EKLAVTVYGVPMKEATTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNENYTE 61
DB 32 EKLAVTVYGVPMKEATTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNENYTE 91
QY 62 HNNMKNMVEQMOEDIIISLMDQSLKPCVKLTPLCGA----- 98
DB 92 HNNMKNMVEQMOEDIIISLMDQSLKPCVKLTPLCGA----- 151
QY 99 -----GCDTSVITQACP 110
DB 152 MMEKEGKNCSEFNISTIRKGYQKEVAFYKLDIIPIDNDTTSYKLTSCNTSVITQACP 211
QY 111 KISFEPIPIHYCAPAGFALIKCNDKTFNGKGPCKKNSTVQCTHGIRPVVSTOLLNLSLA 170
DB 212 KVSFEPPIPIHYCAPAGFALIKCNDKTFNGKGPCKKNSTVQCTHGIRPVVSTOLLNLSLA 271
QY 171 EEEVVRISDGFNNMKTIIIVQLKESVEINCTRPNNNTRKSIH--GPGAFYTTGEIIGD 228
DB 272 EEEVVRISDGFNNMKTIIIVQLKESVEINCTRPNNNTRKSIH--GPGAFYTTGEIIGD 330
QY 229 IROAHCNISRAKNDTLKOIVILKEQF--ENKTIIVFNHSSGGDPEIIVMHSFNGCGEFPYC 287
DB 331 MROAHCNISRAKNDTLKOIVILKEQF--ENKTIIVFNHSSGGDPEIIVMHSFNGCGEFPYC 390
QY 288 NSTQLFNSTWN--NTEGSNTEG--NTITLPCRIOIIMMOEKGAMAPPIRQIRCS 343
DB 391 NSTQLFNSTWN--NTEGSNTEG--NTITLPCRIOIIMMOEKGAMAPPIRQIRCS 450
QY 344 SSNTGLLITRDGGINENGEIIFRPGGDMRDNMSELYKYKVKIEPLGVAFTKCRV 403
DB 451 SSNTGLLITRDGGINENGEIIFRPGGDMRDNMSELYKYKVKIEPLGVAFTKCRV 510
QY 404 VOREGRAVGIGAVFLGAGSTMGASMTLTVQARLLSGIVQOONNLLRAIEAQRM 463
DB 511 VOREGRAVGIGAVFLGAGSTMGASMTLTVQARLLSGIVQOONNLLRAIEAQRM 570
QY 464 LQLTWGIKQLOARVLAVERYLADQDLGIGWCSGKLICTTAVPNNASNSKSLDRIMNN 523
DB 571 LQLTWGIKQLOARVLAVERYLADQDLGIGWCSGKLICTTAVPNNASNSKSLDRIMNN 630
QY 524 MTWMEWESEIDNTSEIYTLIEESONQOEKNEBELLELDKMASLNMWPDITNMLWY 579
DB 631 MTWMEWESEIDNTSEIYTLIEESONQOEKNEBELLELDKMASLNMWPDITNMLWY 686

```

RESULT 12
 US-09-476-242-1
 Sequence 1, Application US/09476242
 Patent No. US2002014663A1
 GENERAL INFORMATION:
 APPLICANT: BARNETT, Susan
 APPLICANT: HARTOG, Karin
 APPLICANT: MARTIN, Eric
 TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES
 FILE REFERENCE: 1605.002

CURRENT APPLICATION NUMBER: US/09/476,242
 CURRENT FILING DATE: 1999-12-30
 NUMBER OF SEQ ID NOS: 26
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO: 1
 LENGTH: 856
 TYPE: PRT
 ORGANISM: Human immunodeficiency virus
 US-09-476-242-1

Query Match 86.4%; Score 2702; DB 10; Length 856;
 Best Local Similarity 79.6%; Pred. No. 3e-246;
 Matches 518; Conservative 28; Mismatches 31; Indels 74; Gaps 6;

```

QY 2 EKLAVTVYGVPMKEATTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNENYTE 61
DB 32 EKLAVTVYGVPMKEATTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNENYTE 91
QY 62 HNNMKNMVEQMOEDIIISLMDQSLKPCVKLTPLCGA----- 98
DB 92 HNNMKNMVEQMOEDIIISLMDQSLKPCVKLTPLCGA----- 151
QY 99 -----GCDTSVITQACPISFE 115
DB 152 GEIKNCSEFNISTIRKGYQKEVAFYKLDIIPIDNDTTSYKLTSCNTSVITQACPISFE 211
QY 116 PPIHYCAPAGFALIKCNDKTFNGKGPCKKNSTVQCTHGIRPVVSTOLLNLSLAEEV 175
DB 212 PPIHYCAPAGFALIKCNDKTFNGKGPCKKNSTVQCTHGIRPVVSTOLLNLSLAEEV 271
QY 176 IRSDFNNAKTIIIVQLKESVEINCTRPNNNTRKSIH--GPGAFYTTGEIIGD 233
DB 272 IRSDFNNAKTIIIVQLKESVEINCTRPNNNTRKSIH--GPGAFYTTGEIIGD 330
QY 234 CNISRAKNDTLKOIVILKEQF--ENKTIIVFNHSSGGDPEIIVMHSFNGCGEFPYCSTOL 292
DB 331 CNISRAKNDTLKOIVILKEQF--ENKTIIVFNHSSGGDPEIIVMHSFNGCGEFPYCSTOL 390
QY 293 FNSTNN--NTEGSNTEG--NTITLPCRIOIIMMOEKGAMAPPIRQIRCS 348
DB 391 FNSTNN--NTEGSNTEG--NTITLPCRIOIIMMOEKGAMAPPIRQIRCS 450
QY 349 GLLITRDGGINENGEIIFRPGGDMRDNMSELYKYKVKIEPLGVAFTKCRVOREK 408
DB 451 GLLITRDGGINENGEIIFRPGGDMRDNMSELYKYKVKIEPLGVAFTKCRVOREK 510
QY 409 RAVGIGAVFLGAGSTMGASMTLTVQARLLSGIVQOONNLLRAIEAQRM 468
DB 511 RAVGIGAVFLGAGSTMGASMTLTVQARLLSGIVQOONNLLRAIEAQRM 570
QY 469 WGIQLOARVLAVERYLADQDLGIGWCSGKLICTTAVPNNASNSKSLDRIMNN 528
DB 571 WGIQLOARVLAVERYLADQDLGIGWCSGKLICTTAVPNNASNSKSLDRIMNN 630
QY 529 WEREIDNTSEIYTLIEESONQOEKNEBELLELDKMASLNMWPDITNMLWY 579
DB 631 WEREIDNTSEIYTLIEESONQOEKNEBELLELDKMASLNMWPDITNMLWY 681

```

RESULT 13
 US-10-196-515-11
 Sequence 11, Application US/10196515
 Publication No. US20030091594A1
 GENERAL INFORMATION:
 APPLICANT: HOXIE, James A.
 APPLICANT: LABRANCHE, Celie C.
 APPLICANT: DOMS, Robert W.
 APPLICANT: HOFFMAN, Trevor L.
 TITLE OF INVENTION: CD4-INDEPENDENT HIV ENVELOPE PROTEINS AS VACCINES AND
 FILE REFERENCE: Hoxie 9596-10401 (0282)
 CURRENT APPLICATION NUMBER: US/10/196,515
 CURRENT FILING DATE: 2002-07-16

PRIOR APPLICATION NUMBER: US/09/337,387
 PRIOR FILING DATE: 1999-06-22
 PRIOR APPLICATION NUMBER: US 09/317,556
 PRIOR FILING DATE: 1999-05-24
 NUMBER OF SEQ ID NOS: 12
 SOFTWARE: Patent Ver. 2.1
 SEQ ID NO 11
 LENGTH: 856
 TYPE: PR
 ORGANISM: Human immunodeficiency virus type 1
 US-10-196-515-11

Query Match 86.1%; Score 2694; DB 15; Length 856;
 Best Local Similarity 79.4%; Pred. No. 1.7e-245;
 Matches 517; Conservative 29; Mismatches 31; Indels 74; Gaps 6;

QY 2 EKLWVTVYGVVWKEATTTLFCASDAKAYDEVHNVWATHACVPTDPNPOEVLENVTE 61
 DB 32 EKLWVTVYGVVWKEATTTLFCASDAKAYDEVHNVWATHACVPTDPNPOEVLENVTE 91
 DB 62 HNNMKNMVBQWQEDIIISLDOSLKPVCVLTPLCGA-----GAG----- 98
 DB 92 NFDNMKNMVBQWQEDIIISLDOSLKPVCVLTPLCGVSLKCTDLKNDNTNTSSGSRMIMEK 151
 QY 99-----GAG-----GAG-----GAG-----GAG-----GAG----- 115
 DB 152 GEIKNCSFNISTGRKVKKEAFYKLDIIPIDNDTSTLSGNTSVITQACPKISFE 211
 QY 116 PIPHYCAPAGFALIKCNDKTFNGKGPCKNVSTVOCTHGIRPVVSTOLLNGSLAEEV 175
 DB 212 PIPHYCAPAGFALIKCNDKTFNGKGPCKNVSTVOCTHGIRPVVSTOLLNGSLAEEV 271
 QY 176 IRSNFTNNAKTIIVOLKESVEINCTRPNNTRKIRIRIQRGPAFTYTK-IGNMROAH 233
 DB 272 IRSNFTNNAKTIIVOLKESVEINCTRPNNTRKIRIRIQRGPAFTYTK-IGNMROAH 330
 QY 234 CNISRAKMNLTLOKIVIKLEQF-ENKTIIVNHSNGSDPEIVMHSFNGGSEFFYCNSITOL 292
 DB 331 CNISRAKMNLTLOKIVIKLEQF-ENKTIIVNHSNGSDPEIVMHSFNGGSEFFYCNSITOL 390
 QY 293 FNSFTNNTG-NTTLPCKRIKOIIMMOEYKAMVAPISQIRCSSNIT 348
 DB 391 FNSFTNNTG-NTTLPCKRIKOIIMMOEYKAMVAPISQIRCSSNIT 450
 QY 349 GLLTRDGINENGTEIFRPGGDMRDMRSELYKVKVLEPLGVAFTKAKRRVOREK 408
 DB 451 GLLTRDGINENGTEIFRPGGDMRDMRSELYKVKVLEPLGVAFTKAKRRVOREK 510
 QY 409 RAVGIGAVFLGFGAAGSTGASMTLTVQARLLSGIVQOONMLRAIEAQORMLQTLV 468
 DB 511 RAVGIGAVFLGFGAAGSTGASMTLTVQARLLSGIVQOONMLRAIEAQORMLQTLV 570
 QY 469 WGIQOLQARVLAVERYLGDQOLGIGWCSGKLICTTAVPNNASMSKSLDRINNMWME 528
 DB 571 WGIQOLQARVLAVERYLGDQOLGIGWCSGKLICTTAVPNNASMSKSLDRINNMWME 630
 QY 529 MEREDINTSELYTLIESONQOEKNEBELLELDKMASLNMWEDITWMLWY 579
 DB 631 MEREDINTSELYTLIESONQOEKNEBELLELDKMASLNMWEDITWMLWY 681

RESULT 14
 US-10-196-515-3
 Sequence 3, Application US/10196515
 Publication No. US20030091594A1
 GENERAL INFORMATION:
 APPLICANT: HOXIE, James A.
 APPLICANT: LABRANCHE, Celja C.
 APPLICANT: DOWS, Robert W.
 APPLICANT: HOFFMAN, Trevor L.
 TITLE OF INVENTION: CD4-INDEPENDENT HIV ENVELOPE PROTEINS AS VACCINES AND
 FILE REFERENCE: Hoxie 9596-104U1 (0282)

CURRENT APPLICATION NUMBER: US/10/196,515
 CURRENT FILING DATE: 2002-07-16
 PRIOR APPLICATION NUMBER: US/09/337,387
 PRIOR FILING DATE: 1999-06-22
 PRIOR APPLICATION NUMBER: US 09/317,556
 PRIOR FILING DATE: 1999-05-24
 NUMBER OF SEQ ID NOS: 12
 SOFTWARE: Patent Ver. 2.1
 SEQ ID NO 3
 LENGTH: 726
 TYPE: PR
 ORGANISM: Human immunodeficiency virus type 1
 US-10-196-515-3

Query Match 85.4%; Score 2672.5; DB 15; Length 726;
 Best Local Similarity 79.0%; Pred. No. 1.4e-243;
 Matches 512; Conservative 31; Mismatches 32; Indels 73; Gaps 7;

QY 2 EKLWVTVYGVVWKEATTTLFCASDAKAYDEVHNVWATHACVPTDPNPOEVLENVTE 61
 DB 32 EKLWVTVYGVVWKEATTTLFCASDAKAYDEVHNVWATHACVPTDPNPOEVLENVTE 91
 DB 62 HNNMKNMVBQWQEDIIISLDOSLKPVCVLTPLCGA-----GAG----- 99
 DB 92 NFDNMKNMVBQWQEDIIISLDOSLKPVCVLTPLCGVSLKCTDLKNDNTNTSSGSRMIMEK 151
 QY 100-----GAG-----GAG-----GAG-----GAG-----GAG----- 115
 DB 152 GEIKNCSFNISTGRKVKKEAFYKLDIIPIDNDTSTLSGNTSVITQACPKISFE 211
 QY 116 PIPHYCAPAGFALIKCNDKTFNGKGPCKNVSTVOCTHGIRPVVSTOLLNGSLAEEV 175
 DB 212 PIPHYCAPAGFALIKCNDKTFNGKGPCKNVSTVOCTHGIRPVVSTOLLNGSLAEEV 271
 QY 176 IRSNFTNNAKTIIVOLKESVEINCTRPNNTRKIRIRIQRGPAFTYTK-IGNMROAH 233
 DB 272 IRSNFTNNAKTIIVOLKESVEINCTRPNNTRKIRIRIQRGPAFTYTK-IGNMROAH 330
 QY 234 CNISRAKMNLTLOKIVIKLEQF-ENKTIIVNHSNGSDPEIVMHSFNGGSEFFYCNSITOL 292
 DB 331 CNISRAKMNLTLOKIVIKLEQF-ENKTIIVNHSNGSDPEIVMHSFNGGSEFFYCNSITOL 390
 QY 293 FNSFTNNTG-NTTLPCKRIKOIIMMOEYKAMVAPISQIRCSSNIT 351
 DB 391 FNSFTNNTG-NTTLPCKRIKOIIMMOEYKAMVAPISQIRCSSNIT 448
 QY 352 LTRPGGINENGTEIFRPGGDMRDMRSELYKVKVLEPLGVAFTKAKRRVOREK 411
 DB 449 LTRPGGINENGTEIFRPGGDMRDMRSELYKVKVLEPLGVAFTKAKRRVOREK 508
 QY 412 GIGAVFLGFGAAGSTGASMTLTVQARLLSGIVQOONMLRAIEAQORMLQTLV 471
 DB 509 GIGAVFLGFGAAGSTGASMTLTVQARLLSGIVQOONMLRAIEAQORMLQTLV 568
 QY 472 KQOLQARVLAVERYLGDQOLGIGWCSGKLICTTAVPNNASMSKSLDRINNMWME 531
 DB 569 KQOLQARVLAVERYLGDQOLGIGWCSGKLICTTAVPNNASMSKSLDRINNMWME 628
 QY 532 EIDNVTSELYTLIESONQOEKNEBELLELDKMASLNMWEDITWMLWY 579
 DB 629 EIDNVTSELYTLIESONQOEKNEBELLELDKMASLNMWEDITWMLWY 676

RESULT 15
 US-10-196-515-12
 Sequence 12, Application US/10196515
 Publication No. US20030091594A1
 GENERAL INFORMATION:
 APPLICANT: HOXIE, James A.
 APPLICANT: LABRANCHE, Celja C.
 APPLICANT: DOWS, Robert W.
 APPLICANT: HOFFMAN, Trevor L.
 TITLE OF INVENTION: CD4-INDEPENDENT HIV ENVELOPE PROTEINS AS VACCINES AND

```

: TITLE OF INVENTION: THERAPEUTICS
: FILE REFERENCE: Hoxie 9596-10411 (0282)
: CURRENT APPLICATION NUMBER: US/10/196,515
: PRIOR FILING DATE: 2002-07-16
: PRIOR APPLICATION NUMBER: US/09/337,387
: PRIOR FILING DATE: 1999-06-22
: PRIOR APPLICATION NUMBER: US 09/317,556
: PRIOR FILING DATE: 1999-05-24
: NUMBER OF SEQ ID NOS: 12
: SOFTWARE: PatentIn Ver. 2.1.
: SEQ ID NO 12
: LENGTH: 759
: TYPE: prt
: ORGANISM: Human immunodeficiency virus type 1
US-10-196-515-12

```

Query Match	84.7%	Score 2651.5;	DB 15;	Length 759;
Best Local Similarity	78.4%;	Pred. No. 1.5e-241;		
Matches 508;	Conservative 32%;	Mismatches 35;	Indels 73;	Gaps 6

```

2  EKXMTVYVVPWPKKATITTLFCASAPAKAYDTEHNWATACVPTDNPCEVLAENTE 61
3  EKXMTVYVVPWPKKATITTLFCASAPAKAYDTEHNWATACVPTDNPCEVLAENTE 61
32 EKXMTVYVVPWPKKATITTLFCASAPAKAYDTEHNWATACVPTDNPCEVLAENTE 91
Qy 62 HFNWKNKNWEOQEDIIISLMDOSLPCVKLTPLCGA----- 98
Db 92 NFNWKNKNWEOQEDIIISLMDOSLPCVKLTPLCGSLKCTDILKNDITNNSGRLIMEK 151
Qy 99 -----GQDSVITIOACPKISFE 115
Db 152 GEIKGCSFNISTSKRGKVKEKAFYKLDIIPIDNDPTSYTLTSCNTSVITIOACPKISFE 211
Qy 116 PIBIHCAPAGFALILKNDKTFENGKPCPKQNVSTVOCTGIRIPVYSTOLLNGSLAESEV 175
Db 212 PIPIHCAPAGFALILKNNKTFNGISPCNNVSTVOCTGIRIPVYSTOLLNGSLAESEV 271
Qy 176 IRSDFNPNNAKTIIVOLKESVBICTRRPNNNTRKSIH--GGRAFYTTGEIICDIPQAH 233
Db 272 IRSNFPNDNAKTIIVOLANTSVBICTCKPNNNTRKSIIRLORGGRAFYTVGK-IGNMGOAH 330
Qy 234 CNISRAKNMTLKOIYIKLREOF-EKKTIVFNHSSGGDBEITMHSFNGCGSEFPYCNSTOL 292
Db 331 CNISRAKNMTLKOIISKLREOFCKNNKTIIFQSSGGBEITMHSFNGCGSEFPYCKSTOL 390
Qy 293 FNSTWNNNTGESSNTEG-NTITLPCRKOIIMMOBVGKAWAPIRGOICRCSNITGL 351
Db 391 FNSTW--STGSSNTEGSDITLPCRKOIIMMOBVGKAWAPIRGOICRCSNITGL 448
Qy 352 LTRDGGINEGTIFPRGGGDRMDNRSELXYKXVYKLEPLGAVPTCKRRVVOREKRAV 411
Db 449 LTRDGGNNNNESEIFPRGGGDRMDNRSELXYKXVYKLEPLGAVPTKRRVVOREKRAV 508
Qy 412 GIGAVFLGFGAAGSTMGAAWMTLTYQALLLSGIVOOONNLRIIEAQQHMLQUTWGI 471
Db 509 GIGALPLGFGAAGSTMGAAWMTLTYQANQISLGIIVOOONNLRIIEQHMLQUTWGI 568
Qy 472 KOLQARVLAVERYLGDOQLLGIWGSCKLICTTAVPANASWNSKSLDIRINNMTWMEER 531
Db 569 KOLQARVLAVERYLKQOQLLGIWGSCKLICTTAVPANASWNSKSLBQINNMTWMEER 628
Qy 532 EIDNNYSEITLIEESONQOEKQOELELDDMAISAMWFDITWMLWY 579
Db 629 EIDNNYSLIHSLEESONQOEKQOELELDDMAISAMWFIISWMLWY 676

```

THIS PAGE BLANK (US:TO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 12:27:29 / Search time 13.4797 Seconds
(without alignments)
4130.780 Million cell updates/sec

Title: US-10-032-162-15

Perfect score: 3129

Sequence: 1 VEKLMVTYGVVPWKAEATL.....ELDKWASLMMNEDITNMLMY 579

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283308 seqs, 96168682 residues

1 number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3043	97.3	847	2 T09448	envelope glycoprot
2	3014	96.3	847	2 S13289	env protein - huma
3	2728.5	87.2	852	2 T12016	envelope glycoprot
4	2724.5	87.1	861	1 VCLJLV	env polypotein pr
5	2719	86.9	856	1 VCLJH3	env polypotein pr
6	2718.5	86.9	856	1 VCLJ3W	env polypotein pr
7	2715	86.8	856	1 VCLJ3V	env polypotein pr
8	2714	86.7	855	1 VCLJ22	env polypotein pr
9	2710	86.6	861	1 VCLJSC	env polypotein pr
10	2707	86.5	854	2 S13288	env protein - huma
11	2706	86.5	843	1 H44001	env polypotein pr
12	2701.5	86.3	851	2 S13985	env polypotein -
13	2696	86.2	852	1 VCLJBR	env polypotein -
14	2679.5	85.6	859	1 VCLJMN	env polypotein pr
15	2628.5	84.0	861	1 VCLJKB	env polypotein pr
16	2628.5	83.9	868	1 VCLJH4	env polypotein -
17	2625.5	83.9	868	1 VCLJH4	env polypotein -
18	2467	78.8	855	1 VCLJZR	env polypotein pr
19	2467	78.1	859	2 S54384	envelope polypote
20	2444.5	78.1	859	2 T01672	envelope polypote
21	2443.5	77.6	856	1 A44963	env polypotein pr
22	2428.5	77.6	846	1 VCLJND	env polypotein pr
23	2060	65.8	854	1 VCLJSI	env polypotein pr
24	2014.5	64.4	443	2 C41621	env polypotein M
25	1997.5	63.8	445	2 A41621	env polypotein M
26	1900	60.7	506	2 A40218	envelop glycoprote
27	1858	59.4	454	2 B41621	env polypotein D
28	1712.5	54.7	495	2 S31493	env polypotein -
29	1688	53.9	863	2 A53034	gag polypotein -

30	1666	53.2	877	2 S49197	envelope protein p
31	1332.5	42.6	290	2 S25940	env protein - huma
32	1331.5	41.9	297	2 S60538	envelope polyprote
33	1213.5	38.8	299	2 S60554	envelope polyprote
34	1208.5	38.6	299	2 S60553	envelope polyprote
35	1203.5	38.5	299	2 S60552	envelope polyprote
36	1199.5	38.3	299	2 S60551	envelope polyprote
37	1198.5	38.3	299	2 S60528	envelope polyprote
38	1198	38.3	294	2 S60545	envelope polyprote
39	1196	38.2	294	2 S60525	envelope polyprote
40	1195.5	38.2	299	2 S60529	envelope polyprote
41	1192.5	38.1	299	2 S60521	envelope polyprote
42	1190.5	38.0	299	2 S60523	envelope polyprote
43	1187.5	38.0	301	2 S60548	envelope polyprote
44	1187	37.9	294	2 S60524	envelope polyprote
45	1182	37.8	300	2 S60546	envelope polyprote

ALIGNMENTS

RESULT 1

T09448

envelope glycoprotein - human immunodeficiency virus type 1 (strain JRFL)

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 24-Nov-1999

C:Accession: T09448

R:Pang, S.; Vinters, H.V.; Akashi, T.; O'Brien, W.A.; Chen, I.S.; Koyanagi, Y.; Namazie,

submitted to the EMBL Data Library, July 1996

A:Reference number: Z1673

A:Accession: T09448

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-847 <PAN>

A:Cross-references: EMBL:U63632; NID:G1465777; PID:G1465781

A:Genetics:

A:Gene: env

C:Superfamily: type E retrovirus env polypotein

Query Match	Score	Length	DB 2;	Length	847;
Best Local Similarity	89.1%	Pred. No. 2.7e-224;			
Matches	573;	Conservative	0;	Mismatches	6;
				Indels	64;
				Gaps	1;
Qy	1	VEKLMVTYGVVPWKAEATTLFCASDAKAVDEVHNMATPACPTDPNPEVLENT	60		
Db	30	VEKLMVTYGVVPWKAEATTLFCASDAKAVDEVHNMATPACPTDPNPEVLENT	89		
Qy	61	EHFNMMKNNMVEQKQEDIIISLMDOSLKPCKVLTPLCGA-----	98		
Db	90	EHFNMMKNNMVEQKQEDIIISLMDOSLKPCKVLTPLCGA-----	149		
Qy	99	-----GCDTSVITQACPKISFEP	116		
Db	150	EIKNCSFNITTSIRDEVOKEYALFYKLDVPIIDNNNTSVRLISCTSVITQACPKISFEP	209		
Qy	117	IPHYCAGAGPAILKCNKNTGNGKPCKNVSTVOCTHGIIPVSVSTOLLNGSLAEVVI	176		
Db	210	IPHYCAGAGPAILKCNKNTGNGKPCKNVSTVOCTHGIIPVSVSTOLLNGSLAEVVI	269		
Qy	177	RSDNFTNNAKTIIVOLKESVEINCTRPNNNTKRSIHIGPRAFYTTGEGIIQDIAHCNI	236		
Db	270	RSDNFTNNAKTIIVOLKESVEINCTRPNNNTKRSIHIGPRAFYTTGEGIIQDIAHCNI	329		
Qy	237	SRAKNDTLKOIVILRQFENKTIIVNHSQGDDEIVNHSFNGCGEPFYCNSTOLFNSI	296		
Db	330	SRAKNDTLKOIVILRQFENKTIIVNHSQGDDEIVNHSFNGCGEPFYCNSTOLFNSI	389		
Qy	297	WNNNTGSSNNTGNTITPCRIKOIINMGEVGMVAPPIRGQIRGCSNITGLLTDG	356		
Db	390	WNNNTGSSNNTGNTITPCRIKOIINMGEVGMVAPPIRGQIRGCSNITGLLTDG	449		
Qy	357	GINENGTEIFRRGGGDMKDNMRSELYKKYKVVYKIEPLGVAPTKCKRNVQREKRAVGIGAV	416		

Db 450 GINENGEIIFRPGGDMRDNRESELYKKYKVKIEPLGVAPTKAKRRVYQREKRAVGICAV 509
Qy 417 FLGLGAGSTMGASMTLTVQARLLLSGIYQOQNNTLRATEAOQRMQLTVWGIKQOA 476
Db 510 FLGLGAGSTMGASMTLTVQARLLLSGIYQOQNNTLRATEAOQRMQLTVWGIKQOA 569
Qy 477 RLVAVERLYGDOQLLGIWCGSGKLICTTAVPNNAWSNKSIDRIINNTMTMEREIDNY 536
Db 570 RLVAVERLYGDOQLLGIWCGSGKLICTTAVPNNAWSNKSIDRIINNTMTMEREIDNY 629
Qy 537 TSEIYTLIEESQNOEKNEOELELDKMASLMMNFDITNMLWY 579
Db 630 TSEIYTLIEESQNOEKNEOELELDKMASLMMNFDITNMLWY 672

RESULT 2

env protein - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997
Accession: S13289
R:O'Brien, W.A.; Koyanagi, Y.; Namazie, A.; Zhao, J.Q.; Diagne, A.; Idler, K.; Zack, J.A.
A:Title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120
A:Reference number: S13288; MUID:91043044; PMID:2172833
A:Accession: S13289
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-847 <OBR>
C:Superfamily: type E retrovirus env polyprotein

Query Match 96.3%; Score 3014; DB 2; Length 847;
Best Local Similarity 88.5%; Pred. No. 4.4e-222;
Matches 568; Conservative 1; Mismatches 9; Indels 64; Gaps 1;

Qy 2 EKLWTVYVYGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENTE 61
Db 31 EKLWTVYVYGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENTE 90
Qy 62 HFNMMKNMVEQWEDIIISLDOSLKPVCVLTPLC-----GAG--- 98
Db 91 HFNMMKNMVEQWEDIIISLDOSLKPVCVLTPLC-----GAG--- 98
Qy 99 -----GAG--- 98
Db 151 IKNSCFNITTRIGNVQKEYALFYKLDVVPIDNTTSYRLINCNTSVITQACPKVSFE 210
Qy 118 PIHYCAPAGFALIKCNDKTFNGKGPCKNVSTVQCTHGIRPVVSTOLLNGSLAESEVY 177
Db 211 PIHYCAPAGFALIKCNDKTFNGKGPCKNVSTVQCTHGIRPVVSTOLLNGSLAESEVY 270
Qy 178 SDNFNNNAKTIIVQLKESVEINCTRPNNNTRKSHIGGRAPYTTGIIIGDIRAHCN 237
Db 271 SDNFNNNAKTIIVQLKESVEINCTRPNNNTRKSHIGGRAPYTTGIIIGDIRAHCN 330
Qy 238 RAKMNDTLKOIYIKAREQENKTIYFNHSSGGDPEIWMHSGGEFFYCNTKLFNS 297
Db 331 RAKMNDTLKOIYIKAREQENKTIYFNHSSGGDPEIWMHSGGEFFYCNTKLFNS 390
Qy 298 NNNTGSGNNTGNTITLPCRKOIINMVEGKAMYPAPIRGOIRCSSNITGLL 357
Db 391 NNNTGSGNNTGNTITLPCRKOIINMVEGKAMYPAPIRGOIRCSSNITGLL 450
Qy 358 INENGTETFRPGGDMRDNRESELYKKYKVKIEPLGVAPTKAKRRVYQREKRAVG 417
Db 451 INENGTETFRPGGDMRDNRESELYKKYKVKIEPLGVAPTKAKRRVYQREKRAVG 510
Qy 418 LGFLGAGSTMGASMTLTVQARLLLSGIYQOQNNTLRATEAOQRMQLTVWGIKQOA 477
Db 511 LGFLGAGSTMGASMTLTVQARLLLSGIYQOQNNTLRATEAOQRMQLTVWGIKQOA 570
Qy 478 VLAVERLYGDOQLLGIWCGSGKLICTTAVPNNAWSNKSIDRIINNTMTMEREIDNY 537
Db 537 VLAVERLYGDOQLLGIWCGSGKLICTTAVPNNAWSNKSIDRIINNTMTMEREIDNY 630

Db 571 VLAVERLYGDOQLLGIWCGSGKLICTTAVPNNAWSNKSIDRIINNTMTMEREIDNY 630
Qy 538 SEIYTLIEESQNOEKNEOELELDKMASLMMNFDITNMLWY 579
Db 631 SEIYTLIEESQNOEKNEOELELDKMASLMMNFDITNMLWY 672

RESULT 3

envelope glycoprotein - human immunodeficiency virus type 1 (strain sc14.3)
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
Accession: T12016
R:McCutchan, F.E.; Sanders-Buell, E.; Salminen, M.O.; Carr, J.K.; Sheppard, W.H.
A:Title: Diversity of the human immunodeficiency virus type 1 envelope glycoprotein in S.
A:Reference number: 217379; MUID:98178716; PMID:9519894
A:Accession: T12016
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-852 <MCC>
A:Cross-references: EMBL:U09034; NID:92351783; PIDN:AACS9271.1; PID:92351784
A:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polyprotein

Query Match 87.2%; Score 2728.5; DB 2; Length 852;
Best Local Similarity 79.9%; Pred. No. 2.9e-200;
Matches 517; Conservative 25; Mismatches 36; Indels 69; Gaps 4;

Qy 2 EKLWTVYVYGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENTE 61
Db 31 EKLWTVYVYGVVWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENTE 90
Qy 62 HFNMMKNMVEQWEDIIISLDOSLKPVCVLTPLC-----GAG--- 99
Db 91 HFNMMKNMVEQWEDIIISLDOSLKPVCVLTPLC-----GAG--- 99
Qy 100 -----GAG--- 99
Db 151 GEINCSFNITTRIGNVQKEYALFYKLDVVPIDNTTSYRLINCNTSVITQACPKVSFE 210
Qy 116 PIHYCAPAGFALIKCNDKTFNGKGPCKNVSTVQCTHGIRPVVSTOLLNGSLAESEVY 175
Db 211 PIHYCAPAGFALIKCNDKTFNGKGPCKNVSTVQCTHGIRPVVSTOLLNGSLAESEVY 270
Qy 176 IRSDNFNNNAKTIIVQLKESVEINCTRPNNNTRKSHIGGRAPYTTGIIIGDIRAHCN 235
Db 271 IRSDNFNNNAKTIIVQLKESVEINCTRPNNNTRKSHIGGRAPYTTGIIIGDIRAHCN 330
Qy 236 ISRAKNDTLKOIYIKAREQENKTIYFNHSSGGDPEIWMHSGGEFFYCNTKLFNS 295
Db 331 ISRAKNDTLKOIYIKAREQENKTIYFNHSSGGDPEIWMHSGGEFFYCNTKLFNS 390
Qy 296 TW--NNNTGSGNNTGNTITLPCRKOIINMVEGKAMYPAPIRGOIRCSSNITGLL 352
Db 391 TW--NNNTGSGNNTGNTITLPCRKOIINMVEGKAMYPAPIRGOIRCSSNITGLL 450
Qy 353 TRDGGINENGTETFRPGGDMRDNRESELYKKYKVKIEPLGVAPTKAKRRVYQREKRAVG 412
Db 451 TRDGGINENGTETFRPGGDMRDNRESELYKKYKVKIEPLGVAPTKAKRRVYQREKRAVG 510
Qy 413 IGALFLGAGSTMGASMTLTVQARLLLSGIYQOQNNTLRATEAOQRMQLTVWGIKQOA 472
Db 511 IGALFLGAGSTMGASMTLTVQARLLLSGIYQOQNNTLRATEAOQRMQLTVWGIKQOA 570
Qy 473 QLOARVLAVERLYGDOQLLGIWCGSGKLICTTAVPNNAWSNKSIDRIINNTMTMERE 532
Db 571 QLOARVLAVERLYGDOQLLGIWCGSGKLICTTAVPNNAWSNKSIDRIINNTMTMERE 630
Qy 533 IDNTSEIYTLIEESQNOEKNEOELELDKMASLMMNFDITNMLWY 579
Db 631 IDNTSEIYTLIEESQNOEKNEOELELDKMASLMMNFDITNMLWY 672

RESULT 4

env polypeptide precursor - human immunodeficiency virus type 1 (isolate LAV-1a)

VCLJLV3

N:Alternate names: coat polypeptide

C:Species: human immunodeficiency virus type 1, HIV-1

A:Note: host Homo sapiens (man)

C:Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 16-Jul-1999

C:Accession: A03975

R:Main-Hobson, S.; Sonigo, P.; Danos, O.; Cole, S.; Allizon, M.

Cell 40, 9-17, 1985

A:Title: Nucleotide sequence of the AIDS virus, LAV.

A:Reference number: A90866; MUID:85099333; PMID:2981635

A:Accession: A03975

A:Molecule type: DNA

A:Residues: 1-861 <MAI>

A:Cross-references: GB:K02013; NID:G326417; PIDN:AAB59751.1; PID:G326424

C:Genetics:

A:Gene: env

A:Superfamily: type E retrovirus env polypeptide

C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyp

F:1-30/Domain: signal sequence #status predicted <SIG>

F:31-516/Product: exterior membrane glycoprotein #status predicted <EXT>

F:517-861/Product: transmembrane glycoprotein #status predicted <TM>

F:186,136,141,146,161,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411

F:616,621,630,642,679,755,821/Binding site: carbohydrate (Asn) (covalent) #status predic

Query Match 87.1%; Score 2724.5; DB 1; Length 861;

Best Local Similarity 79.6%; Pred. No. 66-200;

Matches 522; Conservative 25; Mismatches 30; Indels 79; Gaps 6;

2 EKLWTVYVGVVWKEATTTLFCASDAKAYDEVHANNWATHACVPTDNPQEVLENYTE 61

32 EKLWTVYVGVVWKEATTTLFCASDAKAYDEVHANNWATHACVPTDNPQEVLENYTE 91

62 HNNMKNMVEQMEDIIISLDOSLKPCKLTPLCGA----- 98

92 NFNMMKNDMVEQMEDIIISLDOSLKPCKLTPLCGA----- 151

99 -----GCDTSVITQACP 110

152 MMEKEGKIKNSFNISTIRGKQKQKAYAFYKLDIIPIDNDTSTYLTSCNVSITQACP 211

111 KISFPIPIHYCAPAGFALILKCNKDTFNGKPCPKNVSTVQCTHGIRPVYSTOLLNGSLA 170

212 KVSFPIPIHYCAPAGFALILKCNKDTFNGKPCPKNVSTVQCTHGIRPVYSTOLLNGSLA 271

171 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNTKRSIH--GPRAFYTTGEIGD 228

272 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNTKRSIH--GPRAFYTTGEIGD 228

229 IROAHCNISRAKANDTLKQIVIKLEQF--ENKTIIVFNHSGGDEPIVMSFNGCGEPFYC 287

331 MROAHCNISRAKANDTLKQIVIKLEQF--ENKTIIVFNHSGGDEPIVMSFNGCGEPFYC 390

288 NSTOLFNSTWNN--NTEGSSNNTG--NTITLPCRIKQIINMMQEVGKAMVAPPIRGOIRC 343

391 NSTOLFNSTWNN--NTEGSSNNTG--NTITLPCRIKQIINMMQEVGKAMVAPPIRGOIRC 450

344 SSNTITGLLTRDGGINENGTEIFRPGGDMRDMNRSSELYKYKVKIPELGVAPTKCKRRV 403

451 SSNTITGLLTRDGGINENGTEIFRPGGDMRDMNRSSELYKYKVKIPELGVAPTKCKRRV 510

404 VQREKRAVIGAVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOQNNLLRAIEAQOHL 463

511 VQREKRAVIGAVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOQNNLLRAIEAQOHL 570

464 LQLTWAGIKQOLARLAVERYLGDQQLGIWCGSGKLICTTAVPNNASMSKSLDRIMNN 523

571 LQLTWAGIKQOLARLAVERYLGDQQLGIWCGSGKLICTTAVPNNASMSKSLDRIMNN 630

524 MTEMREIRINDYSEIYTLIEESQNOOEKNEQBLLELDKVASLWNPFDITWLMV 579

RESULT 5

env polypeptide precursor - human immunodeficiency virus type 1 (isolate HTLV-III, BH10)

VCLJLV3

N:Alternate names: coat polypeptide

C:Species: human immunodeficiency virus type 1, HIV-1

A:Note: host Homo sapiens (man)

C:Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 16-Jul-1999

C:Accession: A03973

R:Ratner, L.; Haseltine, W.; Patarca, R.; Lyvak, K.J.; Starcich, B.; Josephs, S.F.; Dora

nberger, J.A.; Papas, T.S.; Ghayee, J.; Chang, N.T.; Gallo, R.C.; Wong-Staal, F.

Nature 313, 277-284, 1985

A:Title: Complete nucleotide sequence of the AIDS virus, HTLV-III.

A:Reference number: A93353; MUID:85111123; PMID:2578615

A:Accession: A03973

A:Molecule type: DNA

A:Residues: 1-856 <RAT>

A:Cross-references: GB:M15654; GB:K02008; GB:K02009; GB:K02010; NID:G326383; PIDN:AAA442

C:Genetics:

A:Gene: env

A:Superfamily: type E retrovirus env polypeptide

C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyp

F:1-30/Domain: signal sequence #status predicted <SIG>

F:31-511/Product: exterior membrane glycoprotein #status predicted <EXT>

F:512-856/Product: transmembrane glycoprotein #status predicted <TM>

F:188,136,141,156,160,186,197,230,234,241,262,276,289,295,301,332,339,356,386,392,397,406

F:611,616,625,637,674,750,816/Binding site: carbohydrate (Asn) (covalent) #status predic

Query Match 86.9%; Score 2719; DB 1; Length 856;

Best Local Similarity 80.0%; Pred. No. 1.66-199;

Matches 521; Conservative 28; Mismatches 28; Indels 74; Gaps 6;

2 EKLWTVYVGVVWKEATTTLFCASDAKAYDEVHANNWATHACVPTDNPQEVLENYTE 61

32 EKLWTVYVGVVWKEATTTLFCASDAKAYDEVHANNWATHACVPTDNPQEVLENYTE 91

62 HNNMKNMVEQMEDIIISLDOSLKPCKLTPLCGA----- 98

92 NFNMMKNDMVEQMEDIIISLDOSLKPCKLTPLCGA----- 151

99 -----GCDTSVITQACP 115

152 GEIKNSCFNISTIRGKQKQKAYAFYKLDIIPIDNDTSTYLTSCNVSITQACP 211

116 PIPHYCAPAGFALILKCNKDTFNGKPCPKNVSTVQCTHGIRPVYSTOLLNGSLAEEBV 175

212 PIPHYCAPAGFALILKCNKDTFNGKPCPKNVSTVQCTHGIRPVYSTOLLNGSLAEEBV 271

176 IRSDNFTNNAKTIIVOLKESVEINCTRPNNNTKRSIH--GPRAFYTTGEIGDIRAH 233

272 IRSDNFTNNAKTIIVOLKESVEINCTRPNNNTKRSIH--GPRAFYTTGEIGDIRAH 233

234 CNISRAKANDTLKQIVIKLEQF--ENKTIIVFNHSGGDEPIVMSFNGCGEPFYCNSIOL 292

331 CNISRAKANDTLKQIVIKLEQF--ENKTIIVFNHSGGDEPIVMSFNGCGEPFYCNSIOL 390

293 FNSITWNN--NTEGSSNNTG--NTITLPCRIKQIINMMQEVGKAMVAPPIRGOIRC 348

391 FNSITWNN--NTEGSSNNTG--NTITLPCRIKQIINMMQEVGKAMVAPPIRGOIRC 450

349 GLILTRDGGINENGTEIFRPGGDMRDMNRSSELYKYKVKIPELGVAPTKCKRRVQREK 408

451 GLILTRDGGINENGTEIFRPGGDMRDMNRSSELYKYKVKIPELGVAPTKCKRRVQREK 510

409 RAVGIGAVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOQNNLLRAIEAQOHLQTLV 468

511 RAVGIGAVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOQNNLLRAIEAQOHLQTLV 570

469 WGIKQOLARLAVERYLGDQQLGIWCGSGKLICTTAVPNNASMSKSLDRIMNNMTME 528

Db 571 WGIKQIARLAVERYLKDQQLLIGWCGSKLICTTAVPWNASWNSKSLQIMNNMTWME 630
Qy 529 MERERIDNTSEIYTLIESONQOEKNEBELLEPKMASLWMPDITWMLWY 579
Db 631 WDRINNTSLIHLIESONQOEKNEBELLEPKMASLWMPDITWMLWY 681

RESULT 6

VCLJ3M

env polypotein precursor - human immunodeficiency virus type 1 (isolate WMJ1)
N:Contains: coat protein gp120; coat protein gp41
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 07-Nov-1997
C:Accession: A24774
R:Starch, B.R.; Hahn, B.H.; Shaw, G.M.; McNeely, P.D.; Modrow, S.; Wolf, H.; Parks, E.
Cell 45, 637-648, 1986
A:Title: Identification and characterization of conserved and variable regions in the env
A:Reference number: A24774; MUID:86218077; PMID:2423250
A:Accession: A24774
A:Molecule type: DNA
A:Residues: 1-856 <STRA>
A:Cross-references: GB:K03455; GB:M38432; NID:G1906382
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polypotein
C:Keywords: coat protein; glycoprotein; polypotein; transmembrane protein
F:1-29/Domin: signal sequence #status predicted <SIG>
F:30-501/Product: coat protein gp120 #status predicted <GP1>
F:502-847/Product: coat protein gp41 #status predicted <GP2>
F:87,134,140,151,155,183,197,234,241,262,276,289,295,331,338,354,360,390,394,404,447,455

Query Match 86.9%; Score 2718.5; DB 1; Length 856;
Best Local Similarity 79.9%; Pred. No. 1-76-199;
Matches 522; Conservative 24; Mismatches 32; Indels 75; Gaps 7;

Qy 1 VEKLVTVYGVVPWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNQEVLNVT 60
Db 30 VEQLMTVYGVVPWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNQEVLNVT 89
Qy 61 EHPNMKNWVEQNEQEDIIISLMDQSLKPCVLTPLC-----GAG-- 99
Db 90 ENFMKNWVEQNEQEDIIISLMDQSLKPCVLTPLC-----GAG-- 149
Qy 100 -----CDTSVITQACPKIS 113
Db 150 KNCSEFNITTSIRKVKHEKAFYKLDIVPIKSNDSITRYRLHCTSVITQACPKIS 209
Qy 114 PEPPIHYCAPAGFALIKCNDKTNGKPCKNVSTVQCTHGIRPVSTOLLNGSLAEE 173
Db 210 PEPPIHYCAPAGFALIKCNDKTNGKPCKNVSTVQCTHGIRPVSTOLLNGSLAEE 269
Qy 174 VVIRSDFTNNAKTIIVOLKESVEINCTRPNNNT-RKSIHTGPRAPYTGELIGDIRQA 232
Db 270 IVIRSENFDAKTIIVHNLNVEINCTRPNNNRRRIHIGPRAPY-TGEIGNRQA 328
Qy 233 HCNISRAKNDTLLKOIYIKLREOFENKTIIVNHSNGGDEIIVNHSNGGDEFFYCNSTOL 292
Db 329 HCNISRAKNDTLLKOIYIKLREOFENKTIIVNHSNGGDEIIVNHSNGGDEFFYCNSTOL 388
Qy 293 FNSFTN--NNTGSSNNTG--GNTITLPCRIKOIINMOEYKAMVAPPIRGQIRCSNI 347
Db 389 FNSFTNVTGISTEGNNTTENGNDITLPCRIKOIINMOEYKAMVAPPIRGQIRCSNI 448
Qy 348 TGLLLTDDGGINENGTEIFRPGGDMRDMNRSSELYKKVVIIEPLGVAFTAKRRVQRE 407
Db 449 TGLLLTDDGSSRSREEIFRPGGDMRDMNRSSELYKKVVIIEPLGVAFTAKRRVQRE 508
Qy 408 KRAVIG-IGAVFLGTGAAGSTGASMTLTVQARLLLSGIYQOQNNLLRAIEAQRMLQ 466
Db 509 KRAVIG-IGAVFLGTGAAGSTGASMTLTVQARLLLSGIYQOQNNLLRAIEAQRMLQ 568
Qy 467 TWVGIKQIARLAVERYLGDQQLLIGWCGSKLICTTAVPWNASWNSKSLQIMNNMTW 526
Db 529 MERERIDNTSEIYTLIESONQOEKNEBELLEPKMASLWMPDITWMLWY 579
Db 631 WDRINNTSLIHLIESONQOEKNEBELLEPKMASLWMPDITWMLWY 681

Db 569 TWVGIKQIARLAVERYLGDQQLLIGWCGSKLICTTAVPWNASWNSKSLQIMNNMTW 628
Qy 527 MERERIDNTSEIYTLIESONQOEKNEBELLEPKMASLWMPDITWMLWY 579
Db 629 MERERIDNTSEIYTLIESONQOEKNEBELLEPKMASLWMPDITWMLWY 681

RESULT 7

VCLJLV

env polypotein precursor - human immunodeficiency virus type 1 (isolate LV)
N:Alternate names: coat polypotein
C:Species: human immunodeficiency virus type 1, HIV-1
A:Note: host Homo sapiens (man)
C:Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 16-Jul-1999
C:Accession: A03974
R:Mesing, M.A.; Smith, D.H.; Cabradilla, C.D.; Benton, C.V.; Lasky, L.A.; Capon, D.J.
Nature 313, 450-458, 1985
A:Title: Nucleic acid structure and expression of the human AIDS lymphadenopathy retrovi
A:Reference number: A93355; MUID:85111157; PMID:2982104
A:Accession: A03974
A:Molecule type: DNA
A:Residues: 1-856 <MBE>
A:Cross-references: GB:K02083; NID:G555008; PIDN:AAB59873.1; PID:G328559
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polypotein
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polypote
F:1-30/Domin: signal sequence #status predicted <SIG>
F:31-511/Product: exterior membrane glycoprotein #status predicted <TM>
F:512-856/Product: transmembrane glycoprotein #status predicted <TM>
F:88,136,141,156,160,186,197,230,234,241,262,276,289,295,301,332,339,356,386,392,397,406,
F:611,616,625,637,674,750,816/Binding site: carbohydrate (Aam) (covalent) #status predict

Query Match 86.8%; Score 2715; DB 1; Length 856;
Best Local Similarity 79.9%; Pred. No. 3-26-199;
Matches 520; Conservative 27; Mismatches 30; Indels 74; Gaps 6;

Qy 2 EKLMTVYGVVPWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNQEVLNVT 61
Db 32 EKLMTVYGVVPWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNQEVLNVT 91
Qy 62 HPMNKNWVEQNEQEDIIISLMDQSLKPCVLTPLCA----- 98
Db 92 NPMNKNWVEQNEQEDIIISLMDQSLKPCVLTPLCA----- 151
Qy 99 -----GCTSVITQACPKISFE 115
Db 152 GEIKNCSFNISTSIKGVQKEVAFYKLDIIPIDNTSYTLTSCNTSVITQACPKISFE 211
Qy 116 PEPPIHYCAPAGFALIKCNDKTNGKPCKNVSTVQCTHGIRPVSTOLLNGSLAEEV 175
Db 212 PEPPIHYCAPAGFALIKCNDKTNGKPCKNVSTVQCTHGIRPVSTOLLNGSLAEEV 271
Qy 176 IRSDFNTNNAKTIIVOLKESVEINCTRPNNNTRKSIH--GPRAPYTGELIGDIRQA 233
Db 272 IRSDFNTNNAKTIIVOLKESVEINCTRPNNNTRKSIH--GPRAPYTGELIGDIRQA 330
Qy 234 CNISRAKNDTLLKOIYIKLREOF-ENKTIIVNHSNGGDEIIVNHSNGGDEFFYCNSTOL 292
Db 331 CNISRAKNDTLLKOIYIKLREOF-ENKTIIVNHSNGGDEIIVNHSNGGDEFFYCNSTOL 390
Qy 293 FNSFTN--NNTGSSNNTG--NNTITLPCRIKOIINMOEYKAMVAPPIRGQIRCSNI 348
Db 391 FNSFTNVTGISTEGNNTTENGNDITLPCRIKOIINMOEYKAMVAPPIRGQIRCSNI 450
Qy 349 GLLLTDDGGINENGTEIFRPGGDMRDMNRSSELYKKVVIIEPLGVAFTAKRRVQRE 408
Db 451 GLLLTDDGGINENGTEIFRPGGDMRDMNRSSELYKKVVIIEPLGVAFTAKRRVQRE 510
Qy 409 RAVGIG-IGAVFLGTGAAGSTGASMTLTVQARLLLSGIYQOQNNLLRAIEAQRMLQ 468
Db 511 RAVGIG-IGAVFLGTGAAGSTGASMTLTVQARLLLSGIYQOQNNLLRAIEAQRMLQ 570

QY 469 WGIKQLQAVLAVERLYGDOQLLIGWCSGSKLICCTAVPWNASWSNKSIDRIINNTWME 528
 Db 571 WGIKQLQAVLAVERLYGDOQLLIGWCSGSKLICCTAVPWNASWSNKSLEQIMNNTWME 630
 QY 528 WEREIDNTSEIYTLIEESQNOQEKNEOBELLELDKASLMMNPDTNNLMY 579
 Db 631 WEREIDNTSEIYTLIEESQNOQEKNEOBELLELDKASLMMNPDTNNLMY 681

RESULT 8

VCLJ32
 env polyprotein precursor - human immunodeficiency virus type 1 (isolate ARV-2)

N:Alternate names: coat polyprotein
 C:Species: human immunodeficiency virus type 1, HIV-1
 A:Note: host Homo sapiens (man)
 C:Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 16-Jul-1999
 C:Accession: A03976
 R:Sanchez-Pescador, R.; Power, M.D.; Barr, P.J.; Steimer, K.S.; Stempien, M.M.; Brown-Sch
 Science 227, 484-492, 1985
 A:Title: Nucleotide sequence and expression of an AIDS-associated retrovirus (ARV-2).
 F:Reference number: A04003; MUID:85090453; PMID:2578227
 C:Accession: A03976

A:Molecule type: DNA

A:Residues: 1-855 <SAN>

A:Cross-references: GB:K02007; MID:G328658; PIDN:AAB59882.1; PID:G328666

A:Genetics:

C:Superfamily: type E retrovirus env polyprotein

C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprot

F:1-30/Domain: signal sequence #status predicted <SIG>

F:31-509/Product: transmembrane glycoprotein #status predicted <TM>

F:510-855/Product: transmembrane glycoprotein #status predicted <TM>

F:87,129,140,158,184,190,200,244,265,282,298,304,334,341,358,364,388,394,400,408,445,458

F:610,624,636,815/binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 86.7%; Score 2714; DB 1; Length 855;

Best Local Similarity 79.3%; Pred. No. 3,8e-199;

Matches 518; Conservative 27; Mismatches 30; Indels 78; Gaps 7;

2 EKLWTVVYGVVWKEATTTLCASDARAYDTEVHNWATACVPTDNPQEVLENTE 61

Db 31 EKLWTVVYGVVWKEATTTLCASDARAYDTEVHNWATACVPTDNPQEVLENTE 90

QY 62 HFNMMKNMNVQMOEDIIISLMDQSLKPCVKLTPLCGA----- 98

Db 91 HFNMMKNMNVQMOEDIIISLMDQSLKPCVKLTPLCGA----- 150

QY 99 -----GCDTSVITQACPK 111

Db 151 EIKNSFNITTSIRDKIQENALFRLLDVVPIDNASTTNTNRYRLIHGNSRVSITQACPK 210

QY 112 ISFEPIPIHYCAPAGPAILKCNKDTNNGKPCKNVSTVOCTHGIRPVSTOLLNGSLAE 171

Db 211 VSFEPIPIHYCTPAGPAILKCNKDTNNGKPCKNVSTVOCTHGIRPVSTOLLNGSLAE 270

QY 172 EEWVIRSDNFTNNAKTIIVQLKESVEINCTRPNNTRKSIHIGPRAFTTGIIIGDIRQ 231

Db 271 EEWVIRSDNFTNNAKTIIVQLKESVEINCTRPNNTRKSIHIGPRAFTTGIIIGDIRQ 330

QY 232 AHCNISRAKNDTLKQIVKLEQF--ENKTIIVFNHSSGGDPEI VMSFNGCFEFCNST 290

Db 331 AHCNISRAKNDTLKQIVKLEQF--ENKTIIVFNHSSGGDPEI VMSFNGCFEFCNST 390

QY 291 QLFNSTNNTNTEGSAANTBGN--TITLPCRIKQIINMOQEVGKAMAPPIRGQIRGSSNT 348

Db 391 QLFNSTNNTNTEGSAANTBGN--TITLPCRIKQIINMOQEVGKAMAPPIRGQIRGSSNT 447

QY 349 GLLLRDGGIN--ENGTEIFRPGGDMRDMRSELKYKVKVKEPIGVATTKKRRVORE 407

Db 448 GLLLRDGGIN--ENGTEIFRPGGDMRDMRSELKYKVKVKEPIGVATTKKRRVORE 507

QY 408 KRAVGI--GAVFLGFLGAAGSTWGAASMTLTVOARLLLSGIYQOONNLLRAIEAOQL 466

Db 508 KRAVGI--GAVFLGFLGAAGSTWGAASMTLTVOARLLLSGIYQOONNLLRAIEAOQL 567
 QY 467 TWGIKQLQAVLAVERLYGDOQLLIGWCSGSKLICCTAVPWNASWSNKSIDRIINNTW 526
 Db 568 TWGIKQLQAVLAVERLYGDOQLLIGWCSGSKLICCTAVPWNASWSNKSLEQIMNNTW 627
 QY 527 WEREIDNTSEIYTLIEESQNOQEKNEOBELLELDKASLMMNPDTNNLMY 579
 Db 628 WEREIDNTSEIYTLIEESQNOQEKNEOBELLELDKASLMMNPDTNNLMY 680

RESULT 9

VCLJ3C
 env polyprotein precursor - human immunodeficiency virus type 1 (isolate SC)

N:Alternate names: coat polyprotein
 C:Species: human immunodeficiency virus type 1, HIV-1
 A:Note: host Homo sapiens (man)
 C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Feb-1997
 C:Accession: B28922
 R:Gurgo, C.; Guo, H.G.; Franchini, G.; Aldovini, A.; Collalti, E.; Farrell, K.; Wong-Sta
 Virology 164, 531-536, 1988
 A:Title: Envelope sequences of two new United States HIV-1 isolates.
 F:Reference number: A28922; MUID:86219542; PMID:3365091
 A:Molecule type: DNA

A:Residues: 1-661 <GUR>

A:Genetics:

C:Superfamily: type E retrovirus env polyprotein

C:Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane prote

F:1-29/Domain: signal sequence #status predicted <SIG>

F:31-861/Product: env polyprotein #status predicted <EP>

F:87,129,135,140,143,159,163,187,198,224,241,262,276,285,301,302,333,340,356,362,386,396

Query Match 86.6%; Score 2710; DB 1; Length 861;

Best Local Similarity 79.2%; Pred. No. 7.7e-199;

Matches 521; Conservative 25; Mismatches 30; Indels 82; Gaps 8;

2 EKLWTVVYGVVWKEATTTLCASDARAYDTEVHNWATACVPTDNPQEVLENTE 61

Db 31 EKLWTVVYGVVWKEATTTLCASDARAYDTEVHNWATACVPTDNPQEVLENTE 90

QY 62 HFNMMKNMNVQMOEDIIISLMDQSLKPCVKLTPLCGA----- 98

Db 91 HFNMMKNMNVQMOEDIIISLMDQSLKPCVKLTPLCGA----- 150

QY 99 -----GCDTSVITQACPK 113

Db 151 KMEGEMTNCSEFNITTSIRSKQKEYALFYKLDVVPIDNNTSYTLINCNSTSVITQACPK 210

QY 114 FEPIPIHYCAPAGPAILKCNKDTNNGKPCKNVSTVOCTHGIRPVSTOLLNGSLAE 173

Db 211 FEPIPIHYCA--RWFALINCNKKNKFTGCTNVSTVOCTHGIRPVSTOLLNGSLAE 269

QY 174 VVIRSDNFTNNAKTIIVQLKESVEINCTRPNNTRKSIHIGPRAFTTGIIIGDIRQ 231

Db 270 VVIRSDNFTNNAKTIIVQLKESVEINCTRPNNTRKSIHIGPRAFTTGIIIGDIRQ 329

QY 232 AHCNISRAKNDTLKQIVKLEQF--ENKTIIVFNHSSGGDPEI VMSFNGCFEFCNST 291

Db 330 AHCNISRAKNDTLKQIVKLEQF--ENKTIIVFNHSSGGDPEI VMSFNGCFEFCNST 389

QY 291 QLFNSTNNTNTEGSAANTBGN--TITLPCRIKQIINMOQEVGKAMAPPIRGQIRGSSNT 347

Db 390 QLFNSTNNTNTEGSAANTBGN--TITLPCRIKQIINMOQEVGKAMAPPIRGQIRGSSNT 448

QY 348 TGLLRDGGINENG-----TEIFRPGGDMRDMRSELKYKVKVKEPIGVATTKKRRVORE 402

Db 449 TGLLRDGGINENG-----TEIFRPGGDMRDMRSELKYKVKVKEPIGVATTKKRRVORE 508

QY 403 VVOREKRAVG--IGAVFLGFLGAAGSTWGAASMTLTVOARLLLSGIYQOONNLLRAIEAO 461

Db 509 VVOREKRAVG--IGAVFLGFLGAAGSTWGAASMTLTVOARLLLSGIYQOONNLLRAIEAO 568

```

QY 462 RMLQLTWGIGKQIOLARVLAVERYLGDQQLIGWCGSGKLICTAVPWNASWNSKSLDRIM 521
D 569 HLLQITWIGIKQIOLARVLAVERYLGDQQLIGWCGSGKLICTTTPVNTSWSNSKSLDRIM 628
QY 522 NMTWMEWEREIDNYTSEIYTLIEESQNOEKNEQELLEDDKMASLWNPEDITWLMWY 579
D 629 GNMWTMEWEREIDNYTSEIYTLIEESQNOEKNEQELLEDDKMASLWNPEDITWLMWY 686

RESULT 10
S13288
env protein - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997
C:Accession: S13288
R:O'Brien, W.A.; Koyanagi, Y.; Namazie, A.; Zhao, J.Q.; Diagne, A.; Idler, K.; Zack, J.A
Nature 348, 69-73, 1990
A:Title: HIV-1 triptem for mononuclear phagocytes can be determined by regions of gp120
Reference number: S13288, MUID:91043044, PMID:2172833
Accession: S13288
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-854 <OBR>
C:Superfamily: type E retrovirus env polyprotein

Query Match 86.5%; Score 2707; DB 2; Length 854;
Best Local Similarity 79.8%; Pred. No. 1.3e-198;
Matches 518; Conservative 25; Mismatches 34; Indels 72; Gaps 6;

QY 2 EKLWVTVYGVVWPKKATTTLFCASDAKAYDEVHNVWATHACVPTDPNQEVLNVTE 61
D 32 EKLWVTVYGVVWPKKATTTLFCASDAKAYDEVHNVWATHACVPTDPNQEVLNVTE 91
QY 62 HFNWKNMMEQOEEDTISLWDSLKPCVKLTPLCGA----- 98
D 92 HFNWKNMMEQOEEDTISLWDSLKPCVKLTPLCGA----- 98
QY 99 -----GCDTSVITQACPKISFEP 117
D 152 GEIKGCSFNITSTRDKVQKAYAFKLDIVPINDTSYRLISCVITQACPKISFEP 211
QY 118 PIHYCAPAGFAILKCNKDFNGKSPCKNVSTVQCTHGIRPVVSTOLLNGSLAEEVIR 177
D 212 PIHYCAPAGFAILKCNKDFNGKSPCKNVSTVQCTHGIRPVVSTOLLNGSLAEEVIR 271
QY 178 SDNTNNAKTIIVOLKESVEINCTRPNNNTKSIHI--GPRRAYTTGELIGDIRAHCHN 235
D 272 SANFTDAKTIIVOLKESVEINCTRPNNNTKSIHI--GPRRAYTTGELIGDIRAHCHN 330
QY 236 ISRAKMMNDTLKOIVIKLREOF--ENKTIYFNHSSGGDEPEIVHSHFNCGEFFYCNSTOLF 294
D 331 ISRAKMMNDTLKOIVIKLREOF--ENKTIYFNHSSGGDEPEIVHSHFNCGEFFYCNSTOLF 390
QY 295 STWNN--NTEGSNNTEG--NTITLPCRKOIINMMEQVGAAPYPIRGOIRCSNITGL 350
D 391 STWNN--NTEGSNNTEG--NTITLPCRKOIINMMEQVGAAPYPIRGOIRCSNITGL 450
QY 351 LITRDGGINENGTEIFRPGGDMRDNRSELKYKVKYKIEPLGVAAPKAKRRVYQREKRA 410
D 451 LITRDGGINENGTEIFRPGGDMRDNRSELKYKVKYKIEPLGVAAPKAKRRVYQREKRA 510
QY 411 VGIGAVFLGFLGAAGSTMGASMTLTVQARLLLSGIYQOQNNLLRAIEAOQHLQLTVWG 470
D 511 VGIGAVFLGFLGAAGSTMGASMTLTVQARLLLSGIYQOQNNLLRAIEAOQHLQLTVWG 570
QY 471 IKOLQAEVLAVERYLGDQQLIGWCGSGKLICTAVPWNASWNSKSLDRIMNNTWME 530
D 571 IKOLQAEVLAVERYLGDQQLIGWCGSGKLICTAVPWNASWNSKSLDRIMNNTWME 630
QY 531 REIDNYTSEIYTLIEESQNOEKNEQELLEDDKMASLWNPEDITWLMWY 579
D 631 REIDNYTSEIYTLIEESQNOEKNEQELLEDDKMASLWNPEDITWLMWY 679

```

```

RESULT 11
H44001
env polyprotein precursor - human immunodeficiency virus type 1 (strain YU-2)
N:Alternate names: coat polyprotein
N:Contains: coat protein gp120; coat protein gp41
C:Species: human immunodeficiency virus type 1, HIV-1
A:Note: host Homo sapiens (man)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Sep-1994
C:Accession: H44001
R:L.I. Y.; Hui, H.; Burgess, C.J.; Price, R.W.; Sharp, P.M.; Hahn, B.H.; Shaw, G.M.
J. Virol. 66, 6587-6600, 1992
A:Title: Complete nucleotide sequence, genome organization, and biological properties of
A:Reference number: A44001; MUID:93021387; PMID:1404605
Accession: H44001
A:Molecule type: DNA
A:Residues: 1-843 <LIY>
A:Cross-references: GB:M93258
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polyprotein
C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein
F:1-29/Domain: signal sequence #status predicted <SIG>
F:19-35/Region: hydrophobic
F:30-489/Product: coat protein gp120 #status predicted <GP1>
F:490-843/Product: coat protein gp41 #status predicted <GP2>
F:499-515/Region: hydrophobic
F:673-689/Region: hydrophobic
F:738-755/Domain: transmembrane #status predicted <TM>
F:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435,
F:437,439,443,446,449,452,455,458,461,464,467,470,473,476,479,482,485,488,491,494,497,500,503,506,509,512,515,518,521,524,527,530,533,536,539,542,545,548,551,554,557,560,563,566,569,572,575,578,581,584,587,590,593,596,599,602,605,608,611,614,617,620,623,626,629,632,635,638,641,644,647,650,653,656,659,662,665,668,671,674,677,680,683,686,689,692,695,698,701,704,707,710,713,716,719,722,725,728,731,734,737,740,743,746,749,752,755,758,761,764,767,770,773,776,779,782,785,788,791,794,797,799,802,805,808,811,814,817,820,823,826,829,832,835,838,841,844,847,850,853,856,859,862,865,868,871,874,877,880,883,886,889,892,895,898,901,904,907,910,913,916,919,922,925,928,931,934,937,940,943,946,949,952,955,958,961,964,967,970,973,976,979,982,985,988,991,994,997,1000,1003,1006,1009,1012,1015,1018,1021,1024,1027,1030,1033,1036,1039,1042,1045,1048,1051,1054,1057,1060,1063,1066,1069,1072,1075,1078,1081,1084,1087,1090,1093,1096,1099,1102,1105,1108,1111,1114,1117,1120,1123,1126,1129,1132,1135,1138,1141,1144,1147,1150,1153,1156,1159,1162,1165,1168,1171,1174,1177,1180,1183,1186,1189,1192,1195,1198,1201,1204,1207,1210,1213,1216,1219,1222,1225,1228,1231,1234,1237,1240,1243,1246,1249,1252,1255,1258,1261,1264,1267,1270,1273,1276,1279,1282,1285,1288,1291,1294,1297,1299,1302,1305,1308,1311,1314,1317,1320,1323,1326,1329,1332,1335,1338,1341,1344,1347,1350,1353,1356,1359,1362,1365,1368,1371,1374,1377,1380,1383,1386,1389,1392,1395,1398,1401,1404,1407,1410,1413,1416,1419,1422,1425,1428,1431,1434,1437,1440,1443,1446,1449,1452,1455,1458,1461,1464,1467,1470,1473,1476,1479,1482,1485,1488,1491,1494,1497,1500,1503,1506,1509,1512,1515,1518,1521,1524,1527,1530,1533,1536,1539,1542,1545,1548,1551,1554,1557,1560,1563,1566,1569,1572,1575,1578,1581,1584,1587,1590,1593,1596,1599,1602,1605,1608,1611,1614,1617,1620,1623,1626,1629,1632,1635,1638,1641,1644,1647,1650,1653,1656,1659,1662,1665,1668,1671,1674,1677,1680,1683,1686,1689,1692,1695,1698,1701,1704,1707,1710,1713,1716,1719,1722,1725,1728,1731,1734,1737,1740,1743,1746,1749,1752,1755,1758,1761,1764,1767,1770,1773,1776,1779,1782,1785,1788,1791,1794,1797,1799,1802,1805,1808,1811,1814,1817,1820,1823,1826,1829,1832,1835,1838,1841,1844,1847,1850,1853,1856,1859,1862,1865,1868,1871,1874,1877,1880,1883,1886,1889,1892,1895,1898,1901,1904,1907,1910,1913,1916,1919,1922,1925,1928,1931,1934,1937,1940,1943,1946,1949,1952,1955,1958,1961,1964,1967,1970,1973,1976,1979,1982,1985,1988,1991,1994,1997,1999,2002,2005,2008,2011,2014,2017,2020,2023,2026,2029,2032,2035,2038,2041,2044,2047,2050,2053,2056,2059,2062,2065,2068,2071,2074,2077,2080,2083,2086,2089,2092,2095,2098,2101,2104,2107,2110,2113,2116,2119,2122,2125,2128,2131,2134,2137,2140,2143,2146,2149,2152,2155,2158,2161,2164,2167,2170,2173,2176,2179,2182,2185,2188,2191,2194,2197,2199,2202,2205,2208,2211,2214,2217,2220,2223,2226,2229,2232,2235,2238,2241,2244,2247,2250,2253,2256,2259,2262,2265,2268,2271,2274,2277,2280,2283,2286,2289,2292,2295,2298,2301,2304,2307,2310,2313,2316,2319,2322,2325,2328,2331,2334,2337,2340,2343,2346,2349,2352,2355,2358,2361,2364,2367,2370,2373,2376,2379,2382,2385,2388,2391,2394,2397,2399,2402,2405,2408,2411,2414,2417,2420,2423,2426,2429,2432,2435,2438,2441,2444,2447,2450,2453,2456,2459,2462,2465,2468,2471,2474,2477,2480,2483,2486,2489,2492,2495,2498,2501,2504,2507,2510,2513,2516,2519,2522,2525,2528,2531,2534,2537,2540,2543,2546,2549,2552,2555,2558,2561,2564,2567,2570,2573,2576,2579,2582,2585,2588,2591,2594,2597,2599,2602,2605,2608,2611,2614,2617,2620,2623,2626,2629,2632,2635,2638,2641,2644,2647,2650,2653,2656,2659,2662,2665,2668,2671,2674,2677,2680,2683,2686,2689,2692,2695,2698,2701,2704,2707,2710,2713,2716,2719,2722,2725,2728,2731,2734,2737,2740,2743,2746,2749,2752,2755,2758,2761,2764,2767,2770,2773,2776,2779,2782,2785,2788,2791,2794,2797,2799,2802,2805,2808,2811,2814,2817,2820,2823,2826,2829,2832,2835,2838,2841,2844,2847,2850,2853,2856,2859,2862,2865,2868,2871,2874,2877,2880,2883,2886,2889,2892,2895,2898,2901,2904,2907,2910,2913,2916,2919,2922,2925,2928,2931,2934,2937,2940,2943,2946,2949,2952,2955,2958,2961,2964,2967,2970,2973,2976,2979,2982,2985,2988,2991,2994,2997,2999,3002,3005,3008,3011,3014,3017,3020,3023,3026,3029,3032,3035,3038,3041,3044,3047,3050,3053,3056,3059,3062,3065,3068,3071,3074,3077,3080,3083,3086,3089,3092,3095,3098,3101,3104,3107,3110,3113,3116,3119,3122,3125,3128,3131,3134,3137,3140,3143,3146,3149,3152,3155,3158,3161,3164,3167,3170,3173,3176,3179,3182,3185,3188,3191,3194,3197,3199,3202,3205,3208,3211,3214,3217,3220,3223,3226,3229,3232,3235,3238,3241,3244,3247,3250,3253,3256,3259,3262,3265,3268,3271,3274,3277,3280,3283,3286,3289,3292,3295,3298,3301,3304,3307,3310,3313,3316,3319,3322,3325,3328,3331,3334,3337,3340,3343,3346,3349,3352,3355,3358,3361,3364,3367,3370,3373,3376,3379,3382,3385,3388,3391,3394,3397,3400,3403,3406,3409,3412,3415,3418,3421,3424,3427,3430,3433,3436,3439,3442,3445,3448,3451,3454,3457,3460,3463,3466,3469,3472,3475,3478,3481,3484,3487,3490,3493,3496,3499,3502,3505,3508,3511,3514,3517,3520,3523,3526,3529,3532,3535,3538,3541,3544,3547,3550,3553,3556,3559,3562,3565,3568,3571,3574,3577,3580,3583,3586,3589,3592,3595,3598,3601,3604,3607,3610,3613,3616,3619,3622,3625,3628,3631,3634,3637,3640,3643,3646,3649,3652,3655,3658,3661,3664,3667,3670,3673,3676,3679,3682,3685,3688,3691,3694,3697,3700,3703,3706,3709,3712,3715,3718,3721,3724,3727,3730,3733,3736,3739,3742,3745,3748,3751,3754,3757,3760,3763,3766,3769,3772,3775,3778,3781,3784,3787,3790,3793,3796,3799,3802,3805,3808,3811,3814,3817,3820,3823,3826,3829,3832,3835,3838,3841,3844,3847,3850,3853,3856,3859,3862,3865,3868,3871,3874,3877,3880,3883,3886,3889,3892,3895,3898,3901,3904,3907,3910,3913,3916,3919,3922,3925,3928,3931,3934,3937,3940,3943,3946,3949,3952,3955,3958,3961,3964,3967,3970,3973,3976,3979,3982,3985,3988,3991,3994,3997,3999,4002,4005,4008,4011,4014,4017,4020,4023,4026,4029,4032,4035,4038,4041,4044,4047,4050,4053,4056,4059,4062,4065,4068,4071,4074,4077,4080,4083,4086,4089,4092,4095,4098,4101,4104,4107,4110,4113,4116,4119,4122,4125,4128,4131,4134,4137,4140,4143,4146,4149,4152,4155,4158,4161,4164,4167,4170,4173,4176,4179,4182,4185,4188,4191,4194,4197,4199,4202,4205,4208,4211,4214,4217,4220,4223,4226,4229,4232,4235,4238,4241,4244,4247,4250,4253,4256,4259,4262,4265,4268,4271,4274,4277,4280,4283,4286,4289,4292,4295,4298,4301,4304,4307,4310,4313,4316,4319,4322,4325,4328,4331,4334,4337,4340,4343,4346,4349,4352,4355,4358,4361,4364,4367,4370,4373,4376,4379,4382,4385,4388,4391,4394,4397,4399,4402,4405,4408,4411,4414,4417,4420,4423,4426,4429,4432,4435,4438,4441,4444,4447,4450,4453,4456,4459,4462,4465,4468,4471,4474,4477,4480,4483,4486,4489,4492,4495,4498,4501,4504,4507,4510,4513,4516,4519,4522,4525,4528,4531,4534,4537,4540,4543,4546,4549,4552,4555,4558,4561,4564,4567,4570,4573,4576,4579,4582,4585,4588,4591,4594,4597,4599,4602,4605,4608,4611,4614,4617,4620,4623,4626,4629,4632,4635,4638,4641,4644,4647,4650,4653,4656,4659,4662,4665,4668,4671,4674,4677,4680,4683,4686,4689,4692,4695,4698,4701,4704,4707,4710,4713,4716,4719,4722,4725,4728,4731,4734,4737,4740,4743,4746,4749,4752,4755,4758,4761,4764,4767,4770,4773,4776,4779,4782,4785,4788,4791,4794,4797,4799,4802,4805,4808,4811,4814,4817,4820,4823,4826,4829,4832,4835,4838,4841,4844,4847,4850,4853,4856,4859,4862,4865,4868,4871,4874,4877,4880,4883,4886,4889,4892,4895,4898,4901,4904,4907,4910,4913,4916,4919,4922,4925,4928,4931,4934,4937,4940,4943,4946,4949,4952,4955,4958,4961,4964,4967,4970,4973,4976,4979,4982,4985,4988,4991,4994,4997,4999,5002,5005,5008,5011,5014,5017,5020,5023,5026,5029,5032,5035,5038,5041,5044,5047,5050,5053,5056,5059,5062,5065,5068,5071,5074,5077,5080,5083,5086,5089,5092,5095,5098,5101,5104,5107,5110,5113,5116,5119,5122,5125,5128,5131,5134,5137,5140,5143,5146,5149,5152,5155,5158,5161,5164,5167,5170,5173,5176,5179,5182,5185,5188,5191,5194,5197,5199,5202,5205,5208,5211,5214,5217,5220,5223,5226,5229,5232,5235,5238,5241,5244,5247,5250,5253,5256,5259,5262,5265,5268,5271,5274,5277,5280,5283,5286,5289,5292,5295,5298,5301,5304,5307,5310,5313,5316,5319,5322,5325,5328,5331,5334,5337,5340,5343,5346,5349,5352,5355,5358,5361,5364,5367,5370,5373,5376,5379,5382,5385,5388,5391,5394,5397,5399,5402,5405,5408,5411,5414,5417,5420,5423,5426,5429,5432,5435,5438,5441,5444,5447,5450,5453,5456,5459,5462,5465,5468,5471,5474,5477,5480,5483,5486,5489,5492,5495,5498,5501,5504,5507,5510,5513,5516,5519,5522,5525,5528,5531,5534,5537,5540,5543,5546,5549,5552,5555,5558,5561,5564,5567,5570,5573,5576,5579,5582,5585,5588,5591,5594,5597,5599,5602,5605,5608,5611,5614,5617,5620,5623,5626,5629,5632,5635,5638,5641,5644,5647,5650,5653,5656,5659,5662,5665,5668,5671,5674,5677,5680,5683,5686,5689,5692,5695,5698,5701,5704,5707,5710,5713,5716,5719,5722,5725,5728,5731,5734,5737,5740,5743,5746,5749,5752,5755,5758,5761,5764,5767,5770,5773,5776,5779,5782,5785,5788,5791,5794,5797,5799,5802,5805,5808,5811,5814,5817,5820,5823,5826,5829,5832,5835,5838,5841,5844,5847,5850,5853,5856,5859,5862,5865,5868,5871,5874,5877,5880,5883,5886,5889,5892,5895,5898,5901,5904,5907,5910,5913,5916,5919,5922,5925,5928,5931,5934,5937,5940,5943,5946,5949,5952,5955,5958,5961,5964,5967,5970,5973,5976,5979,5982,5985,5988,5991,5994,5997,5999,6002,6005,6008,6011,6014,6017,6020,6023,6026,6029,6032,6035,6038,6041,6044,6047,6050,6053,6056,6059,6062,6065,6068,6071,6074,6077,6080,6083,6086,6089,6092,6095,6098,6101,6104,6107,6110,6113,6116,6119,6122,6125,6128,6131,6134,6137,6140,6143,6146,6149,6152,6155,6158,6161,6164,6167,6170,6173,6176,6179,6182,6185,6188,6191,6194,6197,6199,6202,6205,6208,6211,6214,6217,6220,6223,6226,6229,6232,6235,6238,6241,6244,6247,6250,6253,6256,6259,6262,6265,6268,6271,6274,6277,6280,6283,6286,6289,6292,6295,6298,6301,6304,6307,6310,6313,6316,6319,6322,6325,6328,6331,6334,6337,6340,6343,6346,6349,6352,6355,6358,6361,6364,6367,6370,6373,6376,6379,6382,6385,6388,6391,6394,6397,6399,6402,6405,6408,6411,6414,6417,6420,6423,6426,6429,6432,6435,6438,6441,6444,6447,6450,6453,6456,6459,6462,6465,6468,6471,6474,6477,6480,6483,6486,6489,6492,6495,6498,6501,6504,6507,6510,6513,6516,6519,6522,6525,6528,6531,6534,6537,6540,6543,6546,6549,6552,6555,6558,6561,6564,6567,6570,6573,6576,6579,6582,6585,6588,6591,6594,6597,6599,6602,6605,6608,6611,6614,6617,6620,6623,6626,6629,6632,6635,6638,
```

```

Db      567 VLAVERRYLDQDLGIWGSGLICTTVPMTSSNKSINIMWMTMKWERIDNYT 626
Qy      538 SEIYTLIESONQOEKNEBELLELDKMASLMMFDTNNLMY 579
Db      627 HIYSLIESONQOEKNEBELLELDKMASLMMFDTNNLMY 668

```

RESULT 12

S33985

env polyprotein - human immunodeficiency virus type 1

C.Species: human immunodeficiency virus type 1, HIV-1

C.Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 26-Aug-1999

C.Accession: S33985

R.Carlini, F.

submitted to the EMBL Data Library, November 1991

A.Reference number: S33979

A.Accession: S33985

A.Stature: preliminary

A.Molecule type: mRNA

A.Sizes: 1-851 <CAR>

COS-References: EMBL:21530; NID:g60192; PIDN:CAA7628.1; PID:g60199

C.Superfamily: type E retrovirus env polyprotein

Query Match 86.3%; Score 2701.5; DB 2; Length 851;

Best Local Similarity 79.9%; Pred. No. 3.4e-198;

Matches 518; Conservative 27; Mismatches 30; Indels 73; Gaps 6;

```

Qy      2 EKLMTVYVYGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDPNPQVLENTTE 61
Db      32 EKLMTVYVYGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDPNPQVLENTTE 91
Qy      62 HNNMKNMNVBQWEDDIISLWDQSLKPCVKLTPLCA-----GAG----- 98
Db      92 HNNMKNMNVBQWEDDIISLWDQSLKPCVKLTPLCA-----GAG----- 151
Qy      99 -----GDTSVITQACKXISFE 115
Db      152 GEIKNCSFNISTSKRGVQKEYAFYKLDIPIDNDTTSYLTSCNTSVITQACKXISFE 211
Qy      116 PIPHYCAPAGFAILKNDKTFNGKPCKNVSTVOCTGIRPVVSTOLLNGLSLAEEVY 175
Db      212 PIPHYCAPAGFAILKNDKTFNGKPCKNVSTVOCTGIRPVVSTOLLNGLSLAEEVY 271
Qy      176 IRSDFNTNAKTIIVQLKESVEINCTRPNNTRKSIHI--GGRAPYTTGEIIGDIRAH 233
Db      272 IRSDFNTNAKTIIVQLKESVEINCTRPNNTRKSIHI--GGRAPYTTGEIIGDIRAH 330
Qy      234 CNISRAKNDTLKQIVIKLEQF--ENKTIIVFNHSSGGDEPEIWMHSFNGCGEFPYCNSTOL 292
Db      331 CNISRAKNDTLKQIVIKLEQF--ENKTIIVFNHSSGGDEPEIWMHSFNGCGEFPYCNSTOL 390
Qy      293 FNSTNNNTTSGSNNTG--NTITLPCRIRKQIIMMOEVGKAMYPPIRQIRCSSNITGL 351
Db      391 FNST--STKSGSNNTGSDTITLPCRIRKQIIMMOEVGKAMYPPIRQIRCSSNITGL 448
Qy      352 LTRDGIENGTETIRPGGDMRDMRSELVYKVKIPELGVATPKCRVRVQREKAV 411
Db      449 LTRDGIENGTETIRPGGDMRDMRSELVYKVKIPELGVATPKCRVRVQREKAV 508
Qy      412 GIGAVFLGFLGAAGSTMGAASTLTVQARLLLSGIVQOONNLLRAIEAOQRLQLTWGIC 471
Db      509 GIGAVFLGFLGAAGSTMGAASTLTVQARLLLSGIVQOONNLLRAIEAOQRLQLTWGIC 568
Qy      472 KQARVLAVERYLDQDLGIWGSGLICTTAVPWNASMSKSLDRITWNTTWEMER 531
Db      569 KQARVLAVERYLDQDLGIWGSGLICTTAVPWNASMSKSLDRITWNTTWEMER 628
Qy      532 EIDNTSEIYTLIESONQOEKNEBELLELDKMASLMMFDTNNLMY 579
Db      629 EIDNTSEIYTLIESONQOEKNEBELLELDKMASLMMFDTNNLMY 676

```

RESULT 13

VCLJBR

env polyprotein - human immunodeficiency virus type 1 (isolate BR)

N.Alternate names: coat polyprotein

N.Contents: coat protein gp120; coat protein gp41

C.Species: human immunodeficiency virus type 1, HIV-1

A.Note: host Homo sapiens (man)

C.Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 16-Feb-1997

C.Accession: A31667

R.Jand, R. Thayer, R. Srinivasan, A. Nayyar, S. Gardner, M. Luciw, P. Dandekar, S

Virology 168, 79-89, 1989

A.Title: Biological and molecular characterization of human immunodeficiency virus (HIV-

A.Reference number: A94389; MUID:89085613; PMID:2789516

A.Accession: A31667

A.Molecule type: DNA

A.Sizes: 1-852 <ANA>

C.Superfamily: type E retrovirus env polyprotein

C.Keywords: capsid protein; coat protein; polyprotein; transmembrane protein

F.1-516/Product: coat protein gp120 #status predicted <CP1>

F.517-852/Product: coat protein gp41 #status predicted <CP2>

Query Match 86.2%; Score 2696; DB 1; Length 852;

Best Local Similarity 78.7%; Pred. No. 8.9e-198;

Matches 509; Conservative 35; Mismatches 33; Indels 70; Gaps 5;

```

Qy      2 EKLMTVYVYGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDPNPQVLENTTE 61
Db      32 EKLMTVYVYGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDPNPQVLENTTE 91
Qy      62 HNNMKNMNVBQWEDDIISLWDQSLKPCVKLTPLC-----GAG----- 99
Db      92 HNNMKNMNVBQWEDDIISLWDQSLKPCVKLTPLC-----GAG----- 151
Qy      100 -----GDTSVITQACKXISFE 115
Db      152 MKNCSFNITTSIRDMQKEYALFYKLDIPIDNDKTRRYLISGNTSVITQACKXISFE 211
Qy      116 PIPHYCAPAGFAILKNDKTFNGKPCKNVSTVOCTGIRPVVSTOLLNGLSLAEEVY 175
Db      212 PIPHYCAPAGFAILKNDKTFNGKPCKNVSTVOCTGIRPVVSTOLLNGLSLAEEVY 271
Qy      176 IRSDFNTNAKTIIVQLKESVEINCTRPNNTRKSIHI--GGRAPYTTGEIIGDIRAH 235
Db      272 IRSDFNTNAKTIIVQLKESVEINCTRPNNTRKSIHI--GGRAPYTTGEIIGDIRAH 331
Qy      236 ISRAKNDTLKQIVIKLEQF--ENKTIIVFNHSSGGDEPEIWMHSFNGCGEFPYCNSTOL 295
Db      332 ISRAKNDTLKQIVIKLEQF--ENKTIIVFNHSSGGDEPEIWMHSFNGCGEFPYCNSTOL 391
Qy      296 TMNNNTTSGSNNTGNT--ITLPCRIRKQIIMMOEVGKAMYPPIRQIRCSSNITGL 354
Db      392 TMNNNTTSGSNNTGNT--ITLPCRIRKQIIMMOEVGKAMYPPIRQIRCSSNITGL 450
Qy      355 DGGINE--NGTEIFRPGGDMRDMRSELVYKVKIPELGVATPKCRVRVQREKAV 412
Db      451 DGGINE--NGTEIFRPGGDMRDMRSELVYKVKIPELGVATPKCRVRVQREKAV 510
Qy      413 IGAVFLGFLGAAGSTMGAASTLTVQARLLLSGIVQOONNLLRAIEAOQRLQLTWGIC 472
Db      511 IGAVFLGFLGAAGSTMGAASTLTVQARLLLSGIVQOONNLLRAIEAOQRLQLTWGIC 570
Qy      473 QOARVLAVERYLDQDLGIWGSGLICTTAVPWNASMSKSLDRITWNTTWEMER 532
Db      571 QOARVLAVERYLDQDLGIWGSGLICTTAVPWNASMSKSLDRITWNTTWEMER 630
Qy      533 IDNTSEIYTLIESONQOEKNEBELLELDKMASLMMFDTNNLMY 579
Db      631 IDNTSEIYTLIESONQOEKNEBELLELDKMASLMMFDTNNLMY 677

```

RESULT 14

VCLJMN

env polyprotein precursor - human immunodeficiency virus type 1 (isolate MN)

N:Alternate names: coat polyprotein
 C:Species: human immunodeficiency virus type 1, HIV-1
 A:Note: host Homo sapiens (man)
 C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Feb-1997
 C:Accession: A28922
 R:Gargio, C.; Guo, H.G.; Franchini, G.; Aldovini, A.; Collalti, E.; Farrell, K.; Wong-Sta
 Virology 164, 531-536, 1988
 A:Title: Envelope sequences of two new United States HIV-1 isolates.
 A:Reference number: A28922; MUID:88219542; PMID:3369091
 A:Accession: A28922
 A:Molecule type: DNA
 A:Residues: 1-859 <GUR>
 C:Genetics:
 A:Gene: env
 C:Superfamily: type E retrovirus env polyprotein
 C:Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane prote
 F:1-29/Domain: signal sequence #status predicted <SIG>
 F:31-859/Product: env polyprotein #status predicted <EP>
 F:82,129,135,140,141,146,161,165,191,202,246,267,281,294,300,336,343,359,365,389,395,401
 Query Match 85.6%; Score 2679.5; DB 1; Length 859;
 Best Local Similarity 78.2%; Pred. NO. 1.6e-196;
 Matches 513; Conservative 22; Mismatches 42; Indels 79; Gaps 5;

Qy 2 EKLWVTVYGVVPWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLEAVTE 61
 |||||
 Db 31 EKLWVTVYGVVPWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLEAVTE 90
 |||||

Qy 62 HFNNMKNMVEQOMEDIISLMDOSLKPVCVLTPLCGA----- 98
 |||||
 Db 91 HFNNMKNMVEQOMEDIISLMDOSLKPVCVLTPLCVTLNCTDLNNTNNSTANNNSNS 150
 |||||

Qy 99 -----GCDTSVITQAC 109
 |||||
 Db 151 EGTIKGGEKMKNSNITTSIDKQKEVALYKLDIVSINDSTSYRLISCNSTVITQAC 210
 |||||

Qy 110 PKISFEPPIHYCAPAGFAILKNDKTFNGKPCKNSTVQCTHGIRPVYSTOLLNGSL 169
 |||||
 Db 211 PKISFEPPIHYCAPAGFAILKNDKTFNGKPCKNSTVQCTHGIRPVYSTOLLNGSL 270
 |||||

Qy 170 AEEEVVIRSDNFTNNAKTIIVQLKESVEINCTRPNNTRKSIHIGPRAFTTGEIIGDI 229
 |||||
 Db 271 AEEEVVIRSENFNTNAKTIIVHLESVQINCTRPNNTRKSIHIGPRAFTTGTIIGTI 330
 |||||

Qy 230 ROACNISRKAQNDTLKQIVIKLEQEPENKTIIVFNHSSGDPETVMSFNGGEPFNCNS 289
 |||||
 Db 331 ROACNISRKAQNDTLKQIVSKLEQEPENKTIIVFNHSSGDPETVMSFNGGEPFNCNS 390
 |||||

Qy 290 TOLFNSITWNNN---TEGSNNT-EGNTITLPCRKOIIMMOEYKAMAPPIRGQIRCSS 345
 |||||
 Db 391 SPFNSTFWNGNNTWTKGNNTTGSNNNTLQCKIKOIIMMOEYKAMAPPIRGQIRCSS 450
 |||||

Qy 346 NITGLLLTRDGG--INENGTBIIFRPGGDMRDNRSSELYKKYKVKLEPLGVAFTCKRRV 403
 |||||
 Db 451 NITGLLLTRDGGKOTDITDEIIFRPGGDMRDNRSSELYKKYKVKLEPLGVAFTCKRRV 510
 |||||

Qy 404 VQREKRAVIGAVFLGFLGAGSTMGASMTLTVOARLLSGIVQOONNLLRAIEAOORM 463
 |||||
 Db 511 VQREKRA-AIGALFLGFLGAGSTMGASVTLTVQARLLSGIVQOONNLLRAIEAOORM 569
 |||||

Qy 464 LQTLTWGIKQOLAVERYLGDQQLGIWCGSGKLTCTPAVPWNASWSNKSILDRIMNN 523
 |||||
 Db 570 LQTLTWGIKQOLAVERYLGDQQLGIWCGSGKLTCTTVPWNASWSNKSILDRIMNN 629
 |||||

Qy 524 MTWMEEREINVTSEIYTLIEESQNOEKNEOELELDKVASLWNPDTNMTW 579
 |||||
 Db 630 MTWMEEREINVTSEIYTLIEESQNOEKNEOELELDKVASLWNPDTNMTW 685
 |||||

RESULT 15
 VCLAKX
 env polyprotein precursor - human immunodeficiency virus type 1 (strain KB-1-gp32)
 N:Alternate names: coat polyprotein

N:Contains: coat protein gp120; coat protein gp32
 C:Species: human immunodeficiency virus type 1, HIV-1
 A:Note: host Homo sapiens (man)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 01-Mar-1996
 C:Accession: B42995
 R:Shimizu, H.; Hasebe, F.; Tsuchie, H.; Morikawa, S.; Ushijima, H.; Kitamura, T.
 Virology 189, 534-546, 1992
 A:Title: Analysis of a human immunodeficiency virus type 1 isolate carrying a truncated
 A:Reference number: A42995; MUID:92351552; PMID:1322587
 A:Accession: B42995
 A:Molecule type: mRNA
 A:Residues: 1-729 <SHI>
 A:Cross-references: GB:S41266; GB:D01206
 C:Genetics:
 A:Gene: env
 C:Superfamily: type E retrovirus env polyprotein
 C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein
 F:1-689/Domain: extracellular #status predicted <EXT>
 F:1-33/Domain: signal sequence #status predicted <SIG>
 F:17-33/Region: hydrophobic #status predicted
 F:34-517/Product: coat protein gp120 #status predicted <CP1>
 F:514-517/Region: cleavage processing #status predicted <CP2>
 F:518-729/Product: coat protein gp32 #status predicted <CP3>
 F:518-534/Region: hydrophobic #status predicted
 F:690-711/Domain: transmembrane #status predicted <TM1>
 F:712-729/Domain: intracellular #status predicted <INT>
 F:93,141,145,146,163,191,192,237,241,248,269,283,296,308,338,345,361,367,397,403,408,414,
 Query Match 84.0%; Score 2628.5; DB 1; Length 729;
 Best Local Similarity 75.9%; Pred. NO. 1e-192;
 Matches 494; Conservative 33; Mismatches 51; Indels 73; Gaps 2;

Qy 2 EKLWVTVYGVVPWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLEAVTE 61
 |||||
 Db 37 EKLWVTVYGVVPWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLEAVTE 96
 |||||

Qy 62 HFNNMKNMVEQOMEDIISLMDOSLKPVCVLTPLCGA----- 98
 |||||
 Db 97 HFNNMKNMVEQOMEDIISLMDOSLKPVCVLTPLCVTLNCTDLNNTNNNSIEKKIGE 156
 |||||

Qy 99 -----GCDTSVITQACPKIS 113
 |||||
 Db 157 IKONSFVNTINIRDKQKEVALFYKLDVPEIDNNSNTCYRLISCDTSVITQACPKIS 216
 |||||

Qy 114 FEPPIHYCAPAGFAILKNDKTFNGKPCKNSTVQCTHGIRPVYSTOLLNGSLAEE 173
 |||||
 Db 217 FEPPIHYCAPAGFAILKNDKTFNGKPCKNSTVQCTHGIRPVYSTOLLNGSLAEE 276
 |||||

Qy 174 VVIRSDNFTNNAKTIIVQLKESVEINCTRPNNTRKSIHIGPRAFTTGEIIGDIRQAH 233
 |||||
 Db 277 VVIRSENFNTNVTIIIVQLNETYKINCIIRPNNTRKRVMPGRVYVTTGEIIGDIRQAH 336
 |||||

Qy 234 CNISRKAQNDTLKQIVIKLEQEPENKTIIVFNHSSGDPETVMSFNGGEPFNCNSTQLF 293
 |||||
 Db 337 CNISRKAQNDTLKQIVIKLEQEPENKTIIVFNHSSGDPETVMSFNGGEPFNCNSTQLF 396
 |||||

Qy 294 NSTNNNTESNNTSGNTITLPCRKOIIMMOEYKAMAPPIRGQIRCSSNITGLLT 353
 |||||
 Db 397 NSTNLSTGWNGNPENITLPCRKOIIMMOEYKAMAPPIRGQIRCSSNITGLLT 456
 |||||

Qy 354 RDGGINENG---EIFRPGGDMRDNRSSELYKKYKVKLEPLGVAFTCKRRVQREK 408
 |||||
 Db 457 RDGGINNTNNTSIEIFRPGGDMRDNRSSELYKKYKVKLEPLGVAFTCKRRVQREK 516
 |||||

Qy 409 RAVGIGAVFLGFLGAGSTMGASMTLTVOARLLSGIVQOONNLLRAIEAOORMLOLV 468
 |||||
 Db 517 RAVGIGAVFLGFLGAGSTMGAAVTLTVQARQLLPETVQOONNLLRAIDAOHLLQTLV 576
 |||||

Qy 469 WGIKQOLAVERYLGDQQLGIWCGSGKLTCTPAVPWNASWSNKSILDRIMNNMTWME 528
 |||||
 Db 577 WGIKQOLAVERYLGDQQLGIWCGSGKLTCTTVPWNTSNKSFIEINDNTWME 636
 |||||

Qy 529 WEREINDVTSEIYTLIEESQNOEKNEOELELDKVASLWNPDTNMTW 579
 |||||

Db 637 WERINNYTNLYNLIBESONQOEKNEODLALDKWDSLNNWFSITKMLMY 687

Search completed: December 12, 2003, 12:33:33
Job time : 15.4797 secs

THIS PAGE BLANK (US-770)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 12:26:14 ; Search time 29.7807 Seconds
(without alignments)
5017.066 Million cell updates/sec

Title: US-10-032-162-15

Perfect score: 3129
Sequence: 1 VKELMTVYGVPMKEATT.....ELDKNASLWNPDTNLMWY 579

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP rvirus:*
16: SP bacteriap:*
17: SP archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3043	97.3	847	15	Q75760 human immun
2	2824.5	90.3	838	15	Q03806 human immun
3	2814.5	89.9	854	15	Q78225 human immun
4	2800.5	89.5	855	15	Q03805 human immun
5	2789.5	89.1	848	15	Q74999 human immun
6	2784.5	89.0	854	15	Q40222 human immun
7	2780.5	88.9	852	15	Q92761 human immun
8	2777	88.8	853	15	Q03811 human immun
9	2770	88.5	854	15	Q92762 human immun
10	2769	88.3	851	15	Q56110 human immun
11	2764	88.2	863	15	Q77989 human immun
12	2759	88.2	851	15	Q56562 human immun
13	2752.5	88.0	846	15	Q9PXW7 human immun
14	2752	88.0	847	15	Q8AR21 human immun
15	2747	87.8	843	15	Q70150 human immun
16	2745	87.7	861	15	Q9WJVS human immun

17	2743.5	87.7	853	15	O56563 human immun
18	2742	87.6	854	15	O56108 human immun
19	2741.5	87.6	854	15	O92875 human immun
20	2740.5	87.6	843	15	Q70008 human immun
21	2738	87.5	859	15	Q72940 human immun
22	2738	87.5	860	15	Q9E157 human immun
23	2736	87.4	849	15	Q77368 human immun
24	2735.5	87.4	860	15	O9YPS0 human immun
25	2734.5	87.4	868	15	O9MJWS human immun
26	2734	87.4	861	15	O9E527 human immun
27	2733	87.3	861	15	O9E523 human immun
28	2732.5	87.3	841	15	Q9QKJ1 human immun
29	2732	87.3	838	15	O9Q6V0 human immun
30	2730.5	87.3	735	15	O9IWI6 human immun
31	2730.5	87.3	852	15	O56567 human immun
32	2729.5	87.2	856	15	Q72993 human immun
33	2728.5	87.2	852	15	O41883 human immun
34	2726.5	87.1	854	15	O56112 human immun
35	2726	87.1	863	15	O42031 human immun
36	2725	87.1	851	15	Q73304 human immun
37	2724	87.1	857	15	Q8UL70 human immun
38	2723.5	87.0	860	15	Q9DH01 human immun
39	2723	87.0	855	15	Q8Q367 human immun
40	2723	87.0	863	15	Q9WJU8 human immun
41	2722.5	87.0	850	15	Q70003 human immun
42	2722	87.0	752	15	Q70604 human immun
43	2722	87.0	752	15	O70605 human immun
44	2721	87.0	853	15	Q9J023 human immun
45	2720	86.9	851	15	Q8Q852 human immun

ALIGNMENTS

RESULT 1
ID Q75760 PRELIMINARY; PRT; 847 AA.
AC Q75760;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_Taxid=11676;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=JRF;
RA MEDLINE=87206194; PubMed=3646751;
RA Koyanagi Y., Miles S., Mitsuya R.T., Merrill J.E., Vinters H.V.,
Chen I.S.;
RT "Dual infection of the central nervous system by AIDS viruses with
distinct cellular tropisms.";
RL Science 236:819-822(1987).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=JRF;
RA MEDLINE=91043044; PubMed=2172833;
RA O'Brien W.A., Koyanagi Y., Namazie A., Zhao J.Q., Diagne A., Idler K.,
Zack J.A., Chen I.S.;
RT "HIV-1 tropism for mononuclear phagocytes can be determined by regions
of gp120 outside the CD4-binding domain.";
RL Nature 348:69-73(1990).
RN (3)
RP SEQUENCE FROM N.A.
RC STRAIN=JRF;
RA MEDLINE=92092169; PubMed=1684385;
RA Pang S., Vinters H.V., Akashi T., O'Brien W.A., Chen I.S.;
RT "HIV-1 env sequence variation in brain tissue of patients with AIDS-
related neurologic disease.";
RL J. Acquir. Immune Defic. Syndr. 4:1082-1092(1991).
RN (4)

OX NCB1_TaxId=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ba-1;
RX MEDLINE=91289160; PubMed=1905842;
RA Hwang S.S., Boyle T.J., Lyeley H.K., Cullen B.R.;
RT "Identification of the envelope V3 loop as the primary determinant of
RT cell tropism in HIV-1.";
RL Science 253:71-74(1991).
DR EMBL: M63929; AAA75116.1; -.
DR InterPro: IPR000328; Env GP41.
DR InterPro: IPR007777; GP120.
DR Pfam: PF00516; GP120, 1.
DR Pfam: PF00517; GP41, 1.
KW AIDS; Coat protein; Glycoprotein; Polypeptide; Polyprotein; Transmembrane.
SQ SEQUENCE 854 AA; 97186 MW; 005FC881B5C8DFD8 CRC64;

Query Match	89.9%	Score 2814.5;	DB 15;	Length 854;
Best Local Similarity	81.8%	Pred. No. 2,1e-232;		
Matches 531; Conservative	17;	Mismatches 30;	Indels 71;	Gaps 2

Qy	62	HPNNKNNNWEQOQOEIIISLMDPSLPCYKLTPLC-----	96
Dd	31	EKLAMVIVIVCPWAKATITTLFCASAPAKYUVEVHNWATYACVPTDNPQEVLENVTE	61
Dd	91	NPNMNMKNNNWEQOQHEDIIISLMDPSLPCYKLTPLCYLTCTDLDRAATNGDNTNTSSRE	150
Qy	97	--GAG-----CDSVTYQACP	110
Dd	151	MMGGGEMKNCSEFKITTNIRGKQOKEBALYELDIPIDNNSNNRYLISCTSVITYQACP	210
Qy	111	KISFEPPIPIHYCAPAGFALLKCKDNTFNGKGPCKKVS TVQCTHGIRPVSTQLLNGSLA	170
Dd	211	KISFEPPIPIHYCAPAGFALLKCKDKKFNKGPCKSVSTVQCTHGIRPVSTQLLNGSLA	270
Qy	171	EEBVVIRSDNFTNNACTIIIVOLKESEYENCTRPNNNTRKSHIGGRAPFTYTGELIGDIR	230
Dd	271	EEBVVIRSEFAPNACTIIIVOLNESYEINCTRPNNNTRKSHIGGRALYTTGELIGDIR	330
Qy	231	QAHCHNISRAKMNDTLKQIYIKLREOPENKTIYFNHSSGGDPEIWNHSSNCGSEFFYCNST	290
Dd	331	QAHCHNISRAKMNDTLKKIYIKLREOPENKTIYFPGSSGGDEIYHSHNCGSEFFYCNST	390
Qy	291	QLNSTNWNNTBGSNNTBEGNTITLPCRIKOIINMQVGNKAMAPARIGQIRCSSNTTGL	350
Dd	391	QLNSTNWNNTBGSNNTVENNTITLPCRIKOIINMQKGRAMYAPRIGQIRCSSNTTGL	450
Qy	351	LITRDGGINENGTEIFRPGCGDMRDWRSELKYKVYVIEPLGVAFTPCKSKRVVQREKRA	410
Dd	451	LITRDGGEPAKNTKEVFRPGCGDMRDWRSELKYKVYVIEPLGVAFTPCKSKRVVQREKRA	510
Qy	411	VGIGAVFLGFLGAAGSTMGAASHTLTVOARLLLSGIVOOONNLRAIEAQRMQLTVMG	470
Dd	511	VGIGAVFLGFLGAAGSTMGAASHTLTVOARLLLSGIVOOONNLRAIEAQOHLQLTVMG	570
Qy	471	IKOLQARVALAVERYLADDOOLGIMWCSGGLICTTAVPNNASMSNKSIDIRINNNTNME	530
Dd	571	IKOLQARVALAVERYLADDOOLGIMWCSGGLICTTAVPNNASMSNKSINKITDNTNTMEND	630
Qy	531	REIDNTTSELYLIESONQOEKNEOELLEDKMASLNWNTDITNMLAY	579
Dd	631	REINNTTSTIYSLIESONQOEKNEOELLEDKMASLNWNTDITIEWMLY	679

RESULT 4	
003805	
ID 003805	PRELIMINARY;
AC 003805;	PRT; 855 AA.
DT 01-NOV-1996	(TIMESTAMP). 01. Created
DT 01-NOV-1996	(TIMESTAMP). 01. Last sequence update)
DT 01-MAR-2003	(TIMESTAMP). 23. Last amotation update)

DE Env protein.
GN Env.
OS Human immunodeficiency virus 1.
CC Viruses; Retrovirdae; Retroviridae; Lentivirus
OX NCBI_TaxID=11676;
tax

DE Env protein.
GN
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RM [1]
RP SEQUENCE FROM N.A.
RA Reltz M., Popovic M., Gartner S., Gallo R.C., Reed-Connole E.,
RA Beaver B.;
RL Submitted (JUN-1991) to the EMBL/GenBank/DBJ databases.
DR EMBL, M68893; AAA44191.1; --
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
SQ SEQUENCE 855 AA; 96398 MW; EC631A3B84180C8D CRC64;

Query Match	89.5%	Score 2800.5;	DB 15;	Length 855;
Best Local Similarity	81.2%;	Pred. No. 3.3e-231;		
Matches 527;	Conservative 20;	Mismatches 31;	Indels 71;	Gaps 2

Qy	2	EKLMTVVYGVBPWKEATITLFCASDAADAYDEIVANWATACVPTDNPBEVLEANTE	61	
Db	31	EKLMTVVYGVBPWKEATITLFCASDRADYTEVHNWATHACVPTDNPBEVLEAKVTE	90	
Qy	62	HEPMKNNNNVEOMOEIIISLMDOSLKPVCVLTPLC-----	96	
Db	91	NFNMMKNNNNVEOMHEDIIISLMDOSLKPVCVLTPLCVLTCTDLNATNGNDNTTSSRG	150	
Qy	97	--GAG-----	-CPTS VITGACP	110
Db	151	MVGGSGMKRCSFNITITNRKGVKEVAFYKLDIAPIDNNNSNRRLISCTSVITGACP	210	
Qy	111	KISFEEPIPHYCAPAGFALICKNDTTPNGKPCQKSVSTVQCTHGIRPVSTOLLNGSLA	170	
Db	211	KVSFEEPIPHYCAPAGFALICKNDKPKKNGKPCCTNVSTVQCTHGIRPVSTOLLNGSLA	270	
Qy	171	EEBVSIRSDNFTNNAKTTIIIVOLKESVEINCTRPNNNTRKSIIHIGRAFYTTGIIIDIR	230	
Db	271	EEBVSIRSANPDMNAKVIIIVOLNESEVINCTRPNNNTRKSIIHIGRAFYTTGIIIDIR	330	
Qy	231	QAHCHISRAKNMDTLKOIYIKLRBEPENKTI VFNHSSGDEIYTHMSPNCGEFPYCNST	290	
Db	331	QAHCHISRAKNMDTLKIVIKLRBEPGNKTI VFKHSSGDEIYTHSPNCGEFPYCNST	390	
Qy	291	QLENSTMNNNTGESSNTEBNTITLPCR IKOIIINMOWEYGKAMVAPPIRGQIRCSSNITGL	350	
Db	391	QLENSTMNTYESNNTVENNTITLPCR IKOIIINMOWEYGRMVAAPPIRGQIRCSSNITGL	450	
Qy	351	LITROGCIENGTBLFRPGCGDMRPNMSLEYKTKVYKIEBGLGAPFKCRBPVVOPEKRA	410	
Db	451	LITROGCEPDNKTVEFRPGCGDMRPNMSLEYKTKVYKIEBGLGAPFKARBPVVOPEKRA	510	
Qy	411	VGIGAVFLGFLGAAGSTWGAASMTLTVOVARLLLSGIVOOQNNLLRAIEAQQRMLOLTVWG	470	
Db	511	VGIGAVFLGFLGAAGSTWGAAMTLTVOVARLLLSGIVOOQNNLLRAIEAQQHLLDTVWG	570	
Qy	471	IKOLARVLAVERVYGDQDOLGIMWCCSGKLCCTAPVPMNMSWKSXSIDLIMNNMTWEME	530	
Db	571	IKOLARVLAVERVYRDQDOLGIMWCCSGKLCCTAPVPMNMSWKSXSJLNKIMNNMTWEMD	630	
Qy	531	REIDNTYSIIYTLIEESQNOOEKNEQEBLELDKMASIMNWFDTNMLMY	579	
Db	631	REINNTYSIIYGLIEESQNOOEKNEQEBLELDKMASIMNWFDTKMLMY	679	

RESULT 5	ID	PRELIMINARY;	PRT;	848 AA.
074999	Q74999			
AC	Q74999;			
DT	01-NOV-1996	(TRENDArel.	01.	Created)
DT	01-NOV-1996	(TRENDArel.	01.	Last sequence update)
DT	01-OCT-2002	(TRENDArel.	22.	Last annotation update)

```

DE Envelope glycoprotein gp160.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=JR-CSF;
RX MEDLINE=97063128; PubMed=8906996;
RT "Mutations in the vpu, env, and nef genes of a syncytium-inducing
RT variant of HIV type 1 JR-CSF that infects a range of T cell lines."
RL AIDS Res. Hum. Retroviruses 12:347-350(1996).
DR EMBL; U45960; AAB60591.1; -
DR InterPro; IPR000328; Env_Gp41.
DR InterPro; IPR000777; Gp120.
DR Pfam; PF00516; Gp120; 1.
DR Pfam; PF00517; Gp41; 1.
KV AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
FT VARIANT 164 164 N -> D.
FT VARIANT 183 183 S -> N.
FT VARIANT 193 193 N -> S.
FT VARIANT 301 301 I -> T.
FT VARIANT 318 318 K -> E.
FT VARIANT 403 403 Y -> N.
FT VARIANT 525 525 A -> R.
FT VARIANT 833 833 P -> L.
SQ SEQUENCE 848 AA; 96433 MM; B240019C8737BCB3 CRC64;

Query Match 89.1%; Score 2789.5; DB 15; Length 848;
Best Local Similarity 81.9%; Pred. No. 2.8e-230;
Matches 528; Conservative 22; Mismatches 28; Indels 67; Gaps 3;

QY 1 VEKMTVTVYGVVPWKKEATTLTFCASDAKAYDEVHNVMTACVPTDPNQEVLNVT 60
DB 30 VEKMTVTVYGVVPWKKEATTLTFCASDAKAYDEVHNVMTACVPTDPNQEVLNVT 89
QY 61 EHFNMKNMVEQMOEDIIISLMDOSLKPVCVLTPLCGA----- 98
DB 90 EHFNMKNMVEQMOEDIIISLMDOSLKPVCVLTPLCGA----- 149
QY 99 -----GCDTSVITQACPISFE 115
DB 150 EIKNCSFNITSRKRVQKEYALFYKLDVPIIDNNTSYRLINCNTSTITQACPVSFE 209
QY 116 PIPIHYCAPAGFALIKCNDKTPNGKPCKNVSTVOCTHGIRPVSTQLLNGSLAEENV 175
DB 210 PIPIHYCAPAGFALIKCNDKTPNGKPCKNVSTVOCTHGIRPVSTQLLNGSLAEENV 269
QY 176 IRSDFNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGPRAFYTTGELIGDIRQAHN 235
DB 270 IRSDFNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGPRAFYTTGELIGDIRQAHN 329
QY 236 ISRAKNDTLKQIVIKLRQEPENKTIIVNHSFGDPEIWMHSFNGGEPFYCNSQTLFNS 295
DB 330 ISRAKNDTLKQIVIKLRQEPENKTIIVNHSFGDPEIWMHSFNGGEPFYCNSQTLFNS 389
QY 296 TWNNNTESNTEG-NTTLPCRKQIINMOEYKAMYPPIRGQIRCSNITGLLNR 354
DB 390 TW-NDTERKSSGEGVDITILPCRKQIINMOEYKAMYPPIRGQIRCSNITGLLNR 448
QY 355 DGGIENENGTETFRPGGDMRDNRSELKYKVKVKTLEPGVAPTCKRRVOREKAVGIG 414
DB 449 DGGIENENGTETFRPGGDMRDNRSELKYKVKVKTLEPGVAPTCKRRVOREKAVGIG 508
QY 415 AVFLGFLGAAGSTWGAASNTLTVQARLLSGIVQOQNLLRAIEAQOQMLQTLVGIQOL 474
DB 509 ALFLGFLGAAGSTWGAASNTLTVQARLLSGIVQOQNLLRAIEAQOQMLQTLVGIQOL 568
QY 475 QARVLAVERYLGGDQOLLGIWGSGLICTTAVPKNASNSKSLDIANNMTWMEERID 534
DB 569 QARVLAVERYLGGDQOLLGIWGSGLICTTAVPKNASNSKSLDIANNMTWMEERID 628

```

```

QY 535 NYTSEIYTLIEESQNOEKNEOBLELDKWNASLWNPEDITNNLWY 579
DB 629 NYTSEIYTLIEESQNOEKNEOBLELDKWNASLWNPEDITNNLWY 673

RESULT 6
ID 040222 PRELIMINARY; PRT; 854 AA.
AC 040222;
DT 01-JAN-1998 (TREMblrel. 05, Created)
DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
DT 01-OCT-2002 (TREMblrel. 22, Last annotation update)
DE Env polyprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=AD8;
RX MEDLINE=96432129; PubMed=8835195;
RT "Theodore T.S., England G., Buckler-White A., Buckler C.E.,
RT Martin M.A., Peden K.W.;
RT "Construction and characterization of a stable full-length macrophage-
RT tropic HIV type 1 molecular clone that directs the production of high
RT titers of progeny virions."
RL AIDS Res. Hum. Retroviruses 12:191-194(1996).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=AD8;
RA Theodore T.S., England G., Buckler-White A., Buckler C.E.,
RA Martin M.A., Peden K.W.C.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF004394; AAB64170.1; -
DR InterPro; IPR000328; Env_Gp41.
DR InterPro; IPR000777; Gp120.
DR Pfam; PF00516; Gp120; 1.
DR Pfam; PF00517; Gp41; 1.
KV AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 854 AA; 97291 MM; 06C45E691036C12 CRC64;

Query Match 89.0%; Score 2784.5; DB 15; Length 854;
Best Local Similarity 82.0%; Pred. No. 7.7e-230;
Matches 533; Conservative 12; Mismatches 34; Indels 71; Gaps 6;

QY 1 VEKMTVTVYGVVPWKKEATTLTFCASDAKAYDEVHNVMTACVPTDPNQEVLNVT 60
DB 30 VEKMTVTVYGVVPWKKEATTLTFCASDAKAYDEVHNVMTACVPTDPNQEVLNVT 89
QY 61 EHFNMKNMVEQMOEDIIISLMDOSLKPVCVLTPLCGA----- 98
DB 90 EHFNMKNMVEQMOEDIIISLMDOSLKPVCVLTPLCGA----- 149
QY 99 -----GCDTSVITQACPISFE 117
DB 150 EIKNCSFNITSRKRVQKEYALFYKLDVPIIDNNTSYRLINCNTSTITQACPVSFE 209
QY 118 PIPIHYCAPAGFALIKCNDKTPNGKPCKNVSTVOCTHGIRPVSTQLLNGSLAEENV 177
DB 210 PIPIHYCAPAGFALIKCNDKTPNGKPCKNVSTVOCTHGIRPVSTQLLNGSLAEENV 269
QY 210 PIHYCTPAGFALIKCNDKTPNGKPCKNVSTVOCTHGIRPVSTQLLNGSLAEENV 269
DB 178 SDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGPRAFYTTGELIGDIRQAHN 237
DB 270 SDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGPRAFYTTGELIGDIRQAHN 329
QY 238 FAKNNDTLKQIVIKLRQEP-ENKTIIVNHSFGDPEIWMHSFNGGEPFYCNSQTLFNS 296
DB 330 RTKNNTNLNQATLTKQEGFNKTIIVNHSFGDPEIWMHSFNGGEPFYCNSQTLFNS 389
QY 297 WNN-----NTESNTEG-NTTLPCRKQIINMOEYKAMYPPIRGQIRCSNITGL 351
DB 390 WNFNGTWNLTQSNQTEGNDITLPCRKQIINMOEYKAMYPPIRGQIRCSNITGL 449

```

```

QY 352 LTRDGINENG-TEI PRPGGDMRDNRSELYKYKVKIEPLGVAPTKCRVRQREKRA 410
DB 450 LTRDGNHNDTEFRPGGDMRDNRSELYKYKVKIEPLGVAPTKCRVRQREKRA 509
QY 411 VG-ICAVFLGFLGAAGSTMGAASTLTVQARLLSGIVQOONNLLRAIEAOQHLLQLTW 469
DB 510 VGTICAMEFGFGAAGSTMGAASTLTVQARLLSGIVQOONNLLRAIEAOQHLLQLTW 569
QY 470 GIKQOARVLAVERLGRDQQLGIMGCGKLICTTAVPWNASMSKSLDRIMNNNTMWE 529
DB 570 GIKQOARVLAVERLGRDQQLGIMGCGKLICTTAVPWNASMSKSLDRIMNNNTMWE 629
QY 530 EREIDNTSEIYTLIEESQOQEKNEQELBLDKMASLNNPFDITNMLY 579
DB 630 EREIDNTSEIYTLIEESQOQEKNEQELBLDKMASLNNPFDITNMLY 679

RESULT 7
Q92761 PRELIMINARY; PRT; 852 AA.
Q92761:
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=SFMS2.9;
RX MEDLINE=98178716; PubMed=9519894; Salminen M.O., Carr J.K.,
RA McCutchan F.E., Sanders-Buell E., Salminen M.O., Carr J.K.,
RA Sheppard W.H.;
RT "Diversity of the human immunodeficiency virus type 1 envelope
RT glycoprotein in San Francisco Men's Health Study participants.";
RL AIDS Res. Hum. Retroviruses 14:329-337 (1998).
DR EMBL: AF025750; AAC40588.1; -
DR InterPro: IPR000328; Env-CP41.
DR InterPro: IPR000173; GAP-cholesterase.
DR InterPro: IPR000777; GAP-20.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
DR PROSITE: PS00071; GARDH: 1.
KW AIDS; Coat protein; Glycoprotein; Polypeptide; Transmembrane.
SQ SEQUENCE 852 AA; 96545 MW; 3C7780DB0611E617 CRC64;

Query Match 88.9%; Score 2780.5; DB 15; Length 852;
Best Local Similarity 81.1%; Pred. No. 1.7e-229;
Matches 525; Conservative 20; Mismatches 31; Indels 71; Gaps 4;
3 KLVTVVYGVVPWKEATTTLLFCASDAKAYDEVHANNWATHACVPTDPNPOEVLNNT 62
DB 32 QLVTVVYGVVPWKEATTTLLFCASDAKAYDEVHANNWATHACVPTDPNPOEVLNNT 91
QY 63 FNNMKNWVQOQEDIIISLMDOSLKPCVKLTPLCG----- 97
DB 92 FNNMKNWVQOQEDIIISLMDOSLKPCVKLTPLCG----- 151
QY 98 -----AGCDTSVITQACPKISFE 115
DB 152 IKNGSFNITTSIRDKQKDYALFYKLDVVPIDTNNNTSYRLVSCNTSITQACPKISFE 211
QY 116 PIPIHYCAPAGFAILKCKNDKTFNGKPCKNVSTVOCTHGIRPVSTOLLNGLSLAE 175
DB 212 PIPIHYCAPAGFAILKCKNDKTFNGKPCKNVSTVOCTHGIRPVSTOLLNGLSLAE 271
QY 176 IRSDFNTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGRGAFTTGGIITDIOAH 235
DB 272 IRSDFNTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGRGAFTTGGIITDIOAH 331
QY 236 IRSDFNTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGRGAFTTGGIITDIOAH 295

```

```

DB 332 ISRAKMDTLQKQIAQLREQFNKTIIFNOSVGDEPEIVTSHFNGCGEFPYCNASQ 391
QY 236 TNNNTBESNNTGEG--TITLPCRKOIINMVEGKAMYPPIGQIRCSNITGLLT 353
DB 332 TW-NDTBSNNTREDDTITLPCRKOIINMVEGKAMYPPIGQIRCSNITGLLT 450
QY 354 RDGINENG-TEI PRPGGDMRDNRSELYKYKVKIEPLGVAPTKCRVRQREKRA 412
DB 451 RDGINENG-TEI PRPGGDMRDNRSELYKYKVKIEPLGVAPTKCRVRQREKRA 510
QY 413 ICAVFLGFLGAAGSTMGAASTLTVQARLLSGIVQOONNLLRAIEAOQHLLQLTW 472
DB 511 ICAVFLGFLGAAGSTMGAASTLTVQARLLSGIVQOONNLLRAIEAOQHLLQLTW 570
QY 473 QLOARVLAVERLGRDQQLGIMGCGKLICTTAVPWNASMSKSLDRIMNNNTMWE 532
DB 571 QLOARVLAVERLGRDQQLGIMGCGKLICTTAVPWNASMSKSLDRIMNNNTMWE 630
QY 533 IDNTSEIYTLIEESQOQEKNEQELBLDKMASLNNPFDITNMLY 579
DB 631 IDNTSEIYTLIEESQOQEKNEQELBLDKMASLNNPFDITNMLY 677

RESULT 8
Q003811 PRELIMINARY; PRT; 853 AA.
Q003811:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Surface envelope glycoprotein.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=ADA;
RX MEDLINE=91195299; PubMed=2014229;
RA Westervelt P., Gendelman H.E., Ratner L.;
RT "Identification of a determinant within the human immunodeficiency
RT virus 1 surface envelope glycoprotein critical for productive
RT infection of primary monocytes.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:3097-3101 (1991).
DR EMBL: M60472; AAA45065.1; -
DR InterPro: IPR000328; Env-CP41.
DR InterPro: IPR000173; GAP-20.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KW Envelope protein.
SQ SEQUENCE 853 AA; 96998 MW; AD2AF21E2B06AD78 CRC64;

Query Match 88.8%; Score 2777; DB 15; Length 853;
Best Local Similarity 81.4%; Pred. No. 3.4e-229;
Matches 528; Conservative 18; Mismatches 33; Indels 70; Gaps 5;
1 VEKLVTVVYGVVPWKEATTTLLFCASDAKAYDEVHANNWATHACVPTDPNPOEVLNNT 60
DB 30 VENVTVVYGVVPWKEATTTLLFCASDAKAYDEVHANNWATHACVPTDPNPOEVLNNT 89
QY 61 EHENMKNWVQOQEDIIISLMDOSLKPCVKLTPLCGA----- 98
DB 90 EHENMKNWVQOQEDIIISLMDOSLKPCVKLTPLCGA----- 149
QY 99 -----GCDTSVITQACPKISFE 117
DB 150 IKNGSFNITTSIRDKQKDYALFYKLDVVPIDTNNNTSYRLVSCNTSITQACPKISFE 209
QY 118 PIPIHYCAPAGFAILKCKNDKTFNGKPCKNVSTVOCTHGIRPVSTOLLNGLSLAE 177
DB 210 PIPIHYCAPAGFAILKCKNDKTFNGKPCKNVSTVOCTHGIRPVSTOLLNGLSLAE 269
QY 218 SDFNTNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGRGAFTTGGIITDIOAH 237

```



```

Db 270 SSMFTDAKNIIVOLKESVEINCRPNNTKRSIHIGGRAFYTTGELIGDIRQAHCNIS 329
Qy 238 RAKNDLTKQIVLREOF-EKNTIVFNHSGSGPELVNHSFNGCGEPFYCNSTQLNST 296
Db 330 RTKMNNTLNQATLKEQFNGKNTIVFNQSGGPELVNHSFNGCGEPFYCNSTQLNST 389
Qy 297 WNN-----NTGSSNTEGN-TITLPCRQIKOINMMQEVGKAMAPPIRGIRCSNITGL 351
Db 390 WNFNGTNNLQOSNTEGNDITLPCRQIKOINMMQEVGKAMAPPIRGIRCSNITGL 449
Qy 352 LTRDGINENGTETFRPGGDMRDNRSELKYKVKIEPLGVAPTKRRVVOREKRAV 411
Db 450 LTRDGTNNSGSGSEIFRPGGDMRDNRSELKYKVKIEPLGVAPTKRRVVOREKRAV 509
Qy 412 G-IGAVFLGFLGAAGSTMGASMTLVQARLLSGIVQOQNNLLRAIEAQOHLQLTVWG 470
Db 510 GTTGAMFLGFLGAAGSTMGASMTLVQARLLSGIVQOQNNLLRAIEAQOHLQLTVWG 569
Qy 471 IKQOARVLAVERYLGDQQLLGIMGSGKLICTTAVPWNASWSNKSIDRIINNMTWMEWE 530
Db 570 IKQOARVLAVERYLGDQQLLGIMGSGKLICTTAVPWNASWSNKSIDRIINNMTWMEWE 629
Qy 531 REIDNYSSEIYTLIESQNOQEKNEQBELLDKMASLMMNPFDTITNMLWY 579
Db 630 REIDNYSGLIYTLIESQNOQEKNEQBELLDKMASLMMNPFDTITNMLWY 678

```

RESULT 9

```

ID 092762 PRELIMINARY; PRT; 854 AA.
AC 092762;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=SFMSH8.1;
RX MEDLINE=98178716; PubMed=9519894;
RA McCutchan F.E., Sanders-Buell E., Salminen M.O., Carr J.K.,
RA Sheppard W.H.;
RT "Diversity of the human immunodeficiency virus type 1 envelope
RT glycoprotein in San Francisco Men's Health Study participants.";
KM AIDS Res. Hum. Retroviruses 14:329-337(1998).
DR EMBL; AF025754; AAC40591.1; -
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KM AIDS; Coat protein; Glycoprotein; Polypeptide; Transmembrane.
SQ SEQUENCE 854 AA; 97287 MW; AFD8A82B62C5B27D CRC64;

```

```

Query Match 88.5%; Score 2770; DB 15; Length 854;
Best Local Similarity 80.9%; Pred. No. 1,4e-228;
Matches 525; Conservative 23; Mismatches 29; Indels 72; Gaps 5;
Qy 2 EKLMTVYVGVPAWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNTE 61
Db 32 EKLMTVYVGVPAWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNTE 91
Qy 62 HFNMMKNMVEQWQEDIIISLMDQSLKPCVKLTPLCGA----- 98
Db 92 NFNMMKNMVEQWQEDIIISLMDQSLKPCVKLTPLCGA----- 151
Qy 99 -----GCDISVITVQACPKSFE 115
Db 152 GEIKNGSFNITTNRGKMQKEVALFYKLDVVSIDNDNTSYRLISCSNVSITVQACPKVSFO 211

```

```

Qy 116 PIRHYCAPAGFALIKCNDKTENGKPCKNVSTVQCTHGIRPVYSTQLLNGSLAEDEV 175
Db 212 PIRHYCAPAGFALIKCNDKTFSGKPCINVSIVQCTHGIRPVYSTQLLNGSLAEDEV 271
Qy 176 IRSDFNNAKTIIVOLKESVEINCRPNNTKRSIHIGGRAFYTTGELIGDIRQAHCN 235
Db 272 IRSDFNNAKTIIVOLKESVEINCRPNNTKRSIHIGGRAFYTTGELIGDIRQAHCN 331
Qy 236 ISRAKMNDELKQIVLREOFENKTIIVFNHSGSGPELVNHSFNGCGEPFYCNSTQLFNS 295
Db 332 ISRAKMNDELKQIVLREOFENKTIIVFNHSGSGPELVNHSFNGCGEPFYCNSTQLFNS 391
Qy 296 WNN-----NTGSSNTEGN-TITLPCRQIKOINMMQEVGKAMAPPIRGIRCSNITGLLT 353
Db 392 WNFNGTNNLQOSNTEGNDITLPCRQIKOINMMQEVGKAMAPPIRGIRCSNITGLLT 451
Qy 354 RDGINENGT--EIRPGGDMRDNRSELKYKVKIEPLGVAPTKRRVVOREKRA 410
Db 452 RDGGS-NSNGSNKAEVFRPGGDMRDNRSELKYKVKIEPLGVAPTKRRVVOREKRA 510
Qy 411 VGIGAVFLGFLGAAGSTMGASMTLVQARLLSGIVQOQNNLLRAIEAQOHLQLTVWG 470
Db 511 AGIGAVFLGFLGAAGSTMGASMTLVQARLLSGIVQOQNNLLRAIEAQOHLQLTVWG 570
Qy 471 IKQOARVLAVERYLGDQQLLGIMGSGKLICTTAVPWNASWSNKSIDRIINNMTWMEWE 530
Db 571 IKQOARVLAVERYLGDQQLLGIMGSGKLICTTAVPWNASWSNKSIDRIINNMTWMEWE 630
Qy 531 REIDNYSSEIYTLIESQNOQEKNEQBELLDKMASLMMNPFDTITNMLWY 579
Db 631 REIDNYSGLIYTLIESQNOQEKNEQBELLDKMASLMMNPFDTITNMLWY 679

```

RESULT 10

```

ID 056110 PRELIMINARY; PRT; 851 AA.
AC 056110;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=SFMSH8.1;
RX MEDLINE=98178716; PubMed=9519894;
RA McCutchan F.E., Sanders-Buell E., Salminen M.O., Carr J.K.,
RA Sheppard W.H.;
RT "Diversity of the human immunodeficiency virus type 1 envelope
RT glycoprotein in San Francisco Men's Health Study participants.";
KM AIDS Res. Hum. Retroviruses 14:329-337(1998).
DR EMBL; AF025756; AAC40593.1; -
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000173; GAP dhdrogenase.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
DR PROSITE; PS00071; GAPDH; 1.
KM AIDS; Coat protein; Glycoprotein; Polypeptide; Transmembrane.
SQ SEQUENCE 851 AA; 96516 MW; 1CDA83BE6464531 CRC64;

```

```

Query Match 88.5%; Score 2769; DB 15; Length 851;
Best Local Similarity 81.0%; Pred. No. 1,6e-228;
Matches 524; Conservative 22; Mismatches 31; Indels 70; Gaps 3;
Qy 2 EKLMTVYVGVPAWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNTE 61
Db 31 EKLMTVYVGVPAWKEATTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNTE 90
Qy 62 HFNMMKNMVEQWQEDIIISLMDQSLKPCVKLTPLCG----- 97

```

```

Db      91 NNNMMKNNMVEQMEDIISLMDQSLKPCVKLTPLCVTLINCIDLNDNTNTSSSGETMERG 150
Qy      98 -----ACGDSVITQACPKISFEPI 117
Db      151 EIKKNSPITTSIRDKVQKEYALHLKLDVVPIDNDTSYRLVSCNLSVITQACPKVSFEPI 210
Qy      118 PIHYCAPAGFAILKCNKDTFNGKPCKNVSTVOCTHGIRPVVSTOLLNGLSLAESEVIR 177
Db      211 PIHYCAPAGFAILKCNKDRFNGKPCCTNVSTVOCTHGIRPVVSTOPLNGLSLAESEVIR 270
Qy      178 SDNFNNNAKTIIVQLKESVEINCTRPNNNTRKSHIGPGRAYTTGIIIGDIRQACHNIS 237
Db      211 SDNFNNNAKTIIVQLNESVEINCTRPNNNTRKSHIGPGRAYTTGIIIGDIRQACHNIS 330
Qy      238 RAKNMDTLKOIIVIKLREOPENKTIVFNHSSGGDPEIWMHSFNGCGEFPYCNSTOLFNSW 297
Db      331 EAKMHTLBOJAKKREOFKNKTIVFNHSSGGDPEIWMHSFNGCGEFPYCNSTOLFNSW 390
Qy      238 N-----NNTGSSNNTGNTITLPCRICKOIIINMGOEVGKAMVAPIRGOIRCSNITGLL 352
Db      391 SVNSNNTGNTGNNNT-GENITLPCRICKOIIINMGOEVGKAMVAPIRGOIRCSNITGLL 449
Qy      333 TRDGSINENGTETFRPGGDMRDNMRSELYKXKVKIIEPLGVAPTKAKRVRVQREKRAVG 412
Db      450 TRDGSINENGTETFRPGGDMRDNMRSELYKXKVKIIEPLGVAPTKAKRVRVQREKRAVG 509
Qy      413 IGAVLGLGFGAAGSTMGASMTLTVQARLLSGIVQOONNLLRAIEAOQRMLOLTVWGK 472
Db      510 IGAVLGLGFGAAGSTMGASMTLTVQARLLSGIVQOONNLLRAIEAOQRMLOLTVWGK 569
Qy      473 QLOARVLAVERYLGDQQLGIWCGSGKLICTTAVPWNASWSNKSIDRIINNNTMMEWEHE 532
Db      570 QLOARVLAVERYLKQOQLGIWCGSGKLICTTAVPWNASWSNKSIDRIINNNTMMEWEHE 629
Qy      533 IDNTSEIYTLIEESONQOEKNEQELLEDKVASLNNPDTNMLWY 579
Db      630 INNTSLIYTLIEESONQOEKNEQELLEDKVASLNNPDTNMLWY 676

```

RESULT 11

```

Q77989 PRELIMINARY; PRT; 863 AA.
ID Q77989;
AC Q77989;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
Envelope glycoprotein.
ENV.
Human immunodeficiency virus 1.
Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SP128A;
RX MEDLINE=9105685; PubMed=2243391;
RA Liu Z.-Q., Wood C., Levy J.A., Cheng-Mayer C.;
RT "The viral envelope gene is involved in macrophage tropism of a human
RL immunodeficiency virus type 1 strain isolated from brain tissue.";
DR EMBL; M95292; AAA44331.1; -
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120.1.
DR Pfam; PF00517; GP41.1.
KW Aids; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 863 AA; 98222 MW; 0157397FA8B4C5D7 CRC64;

```

```

Query Match      88.3%; Score 2764; DB 15; Length 863;
Best Local Similarity 79.4%; Pred. No. 4, 5e-228;
Matches 523; Conservative 20; Mismatches 34; Indels 82; Gaps 3;

```

Qy 2 EKLWTVVYGVVWKEATTLTFCASDAKAYDTEVHNVNATHACVPTDNPGEVLENTVE 61

```

Db      31 EKLWTVVYGVVWKEATTLTFCASDAKAYDTEVHNVNATHACVPTDNPGEVLENTVE 90
Qy      62 HNNMMKNNMVEQMEDIISLMDQSLKPCVKLTPLCGA----- 98
Db      91 NNNMMKNNMVEQMEDIISLMDQSLKPCVKLTPLCVTLINCIDNLDLNDNTNTDNATSS 150
Qy      99 -----GCDTSVITQAC 109
Db      151 SLRGSTGSIKNSPITTSIRDKVQKEYALFYKLDIVPINNNNTYRLINCNTSITIQAC 210
Qy      110 PKISPEPIPIHYCAPAGFAILKCNKDTFNGKPCKNVSTVOCTHGIRPVVSTOLLNGLSL 169
Db      211 PKVSPEPIPIHYCAPAGFAILKCNKDRFNGKPCKNVSTVOCTHGIRPVVSTOLLNGLSL 270
Qy      170 AEEVVISDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSHIGPGRAYTTGIIIGDI 229
Db      271 AEEVVISDNFTNNAKTIIVQLNESVEINCTRPNNNTRKSHIGPGRAYTTGIIIGDI 330
Qy      230 ROAHGNISRAKNMDTLKOIIVIKLREOPENKTIVFNHSSGGDPEIWMHSFNGCGEFPYCN 289
Db      331 ROAHGTLKQWMDTLKOIIVIKLREOPENKTIVFNHSSGGDPEIWMHSFNGCGEFPYCN 390
Qy      230 TOLFNSWN-----NNTGSSNNTGNTITLPCRICKOIIINMGOEVGKAMVAPIRGO 340
Db      391 TOLFNSWN-----NNTGSSNNTGNTITLPCRICKOIIINMGOEVGKAMVAPIRGO 449
Qy      341 IRCSNITGLLITRDGGINENGTETFRPGGDMRDNMRSELYKXKVKIIEPLGVAPTKCK 400
Db      450 IRCSNITGLLITRDGGINENGTETFRPGGDMRDNMRSELYKXKVKIIEPLGVAPTKCK 509
Qy      401 RRVOREKRAVGIGAVFLGFGAAGSTMGASMTLTVQARLLSGIVQOONNLLRAIEAO 460
Db      510 RRVOREKRAVGIGAVFLGFGAAGSTMGASMTLTVQARLLSGIVQOONNLLRAIEAO 569
Qy      461 QRMLOLTVWGKIQLOARVLAVERYLGDQQLGIWCGSGKLICTTAVPWNASWSNKSIDRI 520
Db      570 QHLLQLTWVGKIQLOARVLAVERYLGDQQLGIWCGSGKLICTTAVPWNASWSNKSIDRI 629
Qy      521 WNNMTMMEWEHEIDNTSEIYTLIEESONQOEKNEQELLEDKVASLNNPDTNMLWY 579
Db      630 WNNMTMMEWEHEIDNTSEIYTLIEESONQOEKNEQELLEDKVASLNNPDTNMLWY 688

```

RESULT 12

```

ID O56562 PRELIMINARY; PRT; 851 AA.
AC O56562;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
Envelope glycoprotein.
ENV.
Human immunodeficiency virus 1.
Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PH1159;
RX MEDLINE=99372987; PubMed=10445815;
RA Ataman-Onal Y., Coiffier C., Giraud A., Babić-Bregg A., Biron F.,
RA Verrier B.;
RT "Comparison of complete env gene sequences from individuals with
RT symptomatic primary HIV type 1 infection.";
RL AIDS Res. Hum. Retroviruses 15:1035-1039(1999).
DR EMBL; AF041128; AAC02519.1; -
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120.1.
DR Pfam; PF00517; GP41.1.
KW Aids; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 851 AA; 96811 MW; 519363B6FBFBF146 CRC64;

```

Query Match 88.2%; Score 2759; DB 15; Length 851;
 Best Local Similarity 80.5%; Pred. No. 1.2e-227;
 Matches 520; Conservative 24; Mismatches 34; Indels 68; Gaps 4;

QY 2 EKLAVTVYGVVWPKKATTTLFCASDAKAYDEVHNVMAHACVPTDPNPOEVLEAVTE 61
 DB 31 EKLAVTVYGVVWPKKATTTLFCASDAKAYDEVHNVMAHACVPTDPNPOEVLEAVTE 90

QY 62 HFNMMKNMVEQWQEDIIISLMDQSLKPCVKLTPLCGA----- 98
 DB 91 HFNMMKNMVEQWQEDIIISLMDQSLKPCVKLTPLCGA----- 150

QY 99 -----GCDTSVITQACPISFPIPIHYCA 123
 DB 151 NYTTIRDKYQKAYALFYKLDVQWDEDNSTYRLISCVTSVITQACPISFPIPIHYCA 210

QY 124 PAGFAILKCNKDTNGKPGCKNVSTVQCTHGIRPVVSTQLLNGSLAEEVVISDNFTN 183
 DB 211 PAGFAILKCNKDTNGKPGCKNVSTVQCTHGIRPVVSTQLLNGSLAEEVVISDNFTN 270

QY 184 NAKTIIVQLKESVEINCTRPNNNTRKSIHIGPRAFYTGEIIGDIRQAHGNISRAKMD 243
 DB 271 NTKTIIVQLKESVEINCTRPNNNTRKSIHIGPRAFYTGEIIGDIRQAHGNISRAKMD 330

QY 244 TLKQIVIKLEQFENKTIIVFNHSSGDEPEIVMHSFNCGGEFFYCNSTOLFNSTWNN-- 300
 DB 331 TLNQIVIKLEQFENKTIIVFNHSSGDEPEIVMHSFNCGGEFFYCNSTOLFNSTWNN-- 390

QY 301 --TEGSNNTEG--NTITLPCRIKOIIMMOEVGKAMAPPIRQOIRCSSNITGLLTRDG 356
 DB 391 NVTRGSNRTGGNTLILPCRIKOIIMMOEVGKAMAPPIRQOIRCSSNITGLLTRDG 450

QY 357 GIN---ENGTEIFRPGGDMRDNRSELYKYKVIKIEPLGVAFTCKRRVVRKRAVGI 413
 DB 451 GNNNNENGTETIFRPGGDMRDNRSELYKYKVIKIEPLGVAFTCKRRVVRKRAVGI 510

QY 414 GAVFLGFGAAGSTGASMTLTVQARLLSGIVQOONLLRAIEAQRMQLDTWGIKQ 473
 DB 511 GAVFLGFGAAGSTGASMTLTVQARLLSGIVQOONLLRAIEAQRMQLDTWGIKQ 570

QY 474 LQAVLAVERYLGGQOLIGTGCSSGLICTAVPNNASWSNKSIDRIANNMTMWEEREI 533
 DB 571 LQAVLAVERYLGGQOLIGTGCSSGLICTAVPNNASWSNKSIDRIANNMTMWEEREI 630

QY 534 DNTSEIYTLIEESQNOQEKNEOELELDKMASLNNWPDITNWLWY 579
 DB 631 DNTSEIYTLIEESQNOQEKNEOELELDKMASLNNWPDITNWLWY 676

RESULT 13
 ID 09PKW7 PRELIMINARY; PRT; 846 AA.
 AC 09PKW7;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
 DE Gp160, envelope glycoprotein.
 GN ENV.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OC NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93189881; PubMed=8446773;
 RA Schmidt-Mayerova H., Gayet O., Guettari N., Bolmont C., Hirsch I.,
 RT "Characterization of HIV-1-PAR, a macrophage-tropic strain: cell
 RT tropism, virus/cell entry and nucleotide sequence of the envelope
 RT glycoprotein."
 RL Res. Virol. 144:21-26(1993).
 DR InterPro: IPR000328; ENV_GP41.
 DR InterPro: IPR000777; GP120.
 DR Pfam: PF00516; GP120; 1.

DR Pfam: PF00517; GP41; 1.
 KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
 SQ SEQUENCE 846 AA; 96490 MW; 3D561D5735F90C84 CRC64;

Query Match 88.0%; Score 2752.5; DB 15; Length 846;
 Best Local Similarity 81.1%; Pred. No. 4.2e-227;
 Matches 522; Conservative 23; Mismatches 30; Indels 69; Gaps 5;

QY 2 EKLAVTVYGVVWPKKATTTLFCASDAKAYDEVHNVMAHACVPTDPNPOEVLEAVTE 61
 DB 31 EKLAVTVYGVVWPKKATTTLFCASDAKAYDEVHNVMAHACVPTDPNPOEVLEAVTE 90

QY 62 HFNMMKNMVEQWQEDIIISLMDQSLKPCVKLTPLCGA----- 98
 DB 91 HFNMMKNMVEQWQEDIIISLMDQSLKPCVKLTPLCGA----- 150

QY 99 -----GCDTSVITQACPISFPIPIHYCA 116
 DB 151 EIKNCSFNITTSIRDKYQKAYALFYKLDVVPIDNSTRYRLISCVTSVITQACPISFPIPIHYCA 210

QY 117 IPIHYCAPAGFAILKCNKDTNGKPGCKNVSTVQCTHGIRPVVSTQLLNGSLAEEVVI 176
 DB 211 IPIHYCAPAGFAILKCNKDTNGKPGCKNVSTVQCTHGIRPVVSTQLLNGSLAEEVVI 270

QY 177 RSDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPRAFYTGEIIGDIRQAHGNT 236
 DB 271 RSDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPRAFYTGEIIGDIRQAHGNT 330

QY 237 SRKANDTLKOIYIKLEQFENKTIIVFNHSSGDEPEIVMHSFNCGGEFFYCNSTOLFNST 296
 DB 331 SKTKMETFQOIVKLEQFENKTIIVFNHSSGDEPEIVMHSFNCGGEFFYCNSTOLFNST 390

QY 297 MNNNTEGSSNTEGNTITLPCRIKOIIMMOEVGKAMAPPIRQOIRCSSNITGLLTRDG 356
 DB 391 W-NDTEGSSNTE-RTITLPCRIKOIIMMOEVGKAMAPPIRQOIRCSSNITGLLTRDG 448

QY 357 G-INENGTEIFRPGGDMRDNRSELYKYKVIKIEPLGVAFTCKRRVVRKRAVGI 415
 DB 449 GNNNNNTEIFRPGGDMRDNRSELYKYKVIKIEPLGVAFTCKRRVVRKRAVGI 508

QY 416 VFGLGFGAAGSTGASMTLTVQARLLSGIVQOONLLRAIEAQRMQLDTWGIKQ 475
 DB 509 VFGLGFGAAGSTGASMTLTVQARLLSGIVQOONLLRAIEAQRMQLDTWGIKQ 568

QY 476 ARVLAVERYLGGQOLIGTGCSSGLICTAVPNNASWSNKSIDRIANNMTMWEEREIDN 535
 DB 569 ARVLAVERYLGGQOLIGTGCSSGLICTAVPNNASWSNKSIDRIANNMTMWEEREIDN 627

QY 536 YTSSEIYTLIEESQNOQEKNEOELELDKMASLNNWPDITNWLWY 579
 DB 628 YTSSEIYTLIEESQNOQEKNEOELELDKMASLNNWPDITNWLWY 671

RESULT 14
 ID 08AR21 PRELIMINARY; PRT; 847 AA.
 AC 08AR21;
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Envelope glycoprotein.
 GN ENV.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OC NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=SHIVSF162P3.2;
 RA Gao F.;
 RL Submitted (Aug-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF536757; AAN05642.1; --
 DR SEQUENCE 847 AA; 95879 MW; 9CF863B97B3BB54C CRC64;

	Query Match	88.0%	Score 2752;	DB 15;	Length 847;
	Best Local Similarity	81.4%;	Pred. No. 4.7e-227;		
	Matches	525;	Conservative 17;	Mismatches 35;	Indels 68; Gaps 4

Qy	1	VEKLVAVTVYGVVPEWKEATTTTLFCASDAAYTEVHNWATHACVTDENPQEVLENVT	60
Dd	30	VEKLVAVTVYGVVPAKKEATTTTLFCASDAQAYDEVHNVNATHACVTDENPQEVLENVT	89
Qy	61	EHEFNWKNNMVEQMEDIIISLDQSILKPCVKLTPLCGA-----	98
Dd	90	EHEFNWKNNMVEQMEDIIISLDQSLPCVKLTPLCVTLHLCTVLHENATWTSSNMKEMNR	149
Qy	99	-----GCSTSYTQAQCPKLSSE	115
Dd	150	GEIKNCSEFVNTTSIGNKKQKEYALFYRLDVPIDNDNTSYNLINCNSTVITQACPVSSE	209
Qy	116	PIPIHYCAPAGFAILKCDNKDTFNGKPCKNVSTVOCTHGIRPVSTQLLNGSLAEENV	175
Dd	210	PIPIHYCAPAGFAILLKCDNKDTFNGSPCLNVSTVOCTHGIRPVSTQLLNGSLAEENV	269
Qy	176	IRSDFNFNNAKTIIVOLKESVEINCTRPNNNTRKSIHIIGGRAFTYTGEIIGDIRAHCH	235
Dd	270	IRSEFNTDNVKTIIVOLKESVEINCTRPNNNTRKSIPIGPGRKFATGDIIGDIRAHCH	329
Qy	226	ISRACKNDTKIQIVIKLRQEPENKTIVPHSHSGSDPEIYMHSNCCGEFFCYNSTOLFNS	295
Dd	330	ISGEKWNTLKQIVKRLQAFENKTIIVFKOSSGDPEIYMHSNCCGEFFCYNSTOLFNS	389
Qy	296	TWNNTGESNNNEGNTITLPCKRIKOIINNMOEGKMVAAPRIGQIRCSNITGLLLTSD	355
Dd	390	TW-ANTIGPNNTNG-TITLPCKRIKOIINNMOEYGKMYAPRIGQIRCSNITGLLLTSD	447
Qy	356	GGIN-ENGTEIPRPGGGMDRWDRSELKYKVKYLEPLGVAPTCKRRVVOREKAHVIG	414
Dd	448	GGREVGNTTEIPRPGGGMDRWDRSELKYKVKYLEPLGVAPTKARRVVQREKRAVTIG	507
Qy	415	AVFLGFLCAGSTMGAASNTLTVQARLLLSGIVOOQNLLRAIEAOQRMLOLTWGIKOL	474
Dd	508	AVFLGFLCAGSTMGAASLTLTVQARLLLSGIVOOQNLLRAIEAOQRLLQJLTWGIKOL	567
Qy	475	OARVLAVERLYGDDOOLLIGWGSGKLICTAAPNMAWSNKSILDRIMNNMTWMEMEREID	534
Dd	568	OARVLAVERYLKDQOLLIGWGSGKLICTTAAPNMAWSNKSILDQIMNNMTWMEMEREIG	627
Qy	535	NYSSEIYTLIESONQOEKNQEELDLDMASIMNFDITNNLMY	579
Dd	628	NYTNLIYTLIESONQOEKNQEELDLDMASIMNMLDISKMLMY	672

CURT 15

ID	070150	PRELIMINARY;	PRT;	843 AA.
AC	070150;			
DT	01-NOV-1996 (TREMBLrel. 01,	Created)		
DT	01-NOV-1996 (TREMBLrel. 01,	Last sequence update)		
DT	01-OCT-2002 (TREMBLrel. 22,	Last annotation update)		
Dd	Envelope glycoprotein (Fragment).			
GN	ENV.			
OS	Human immunodeficiency virus 1.			
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.			
CX	NCBI_TaxId=11676;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=014;			
RX	MEDLINE=95194694; PubMed=7888189;			
RA	Gao F., Yue L., Craig S., Thornton C.L., Robertson D.L.,			
RA	McClellan F.B., Brade J.A., Sharp P.M., Hahn B.H.;			
RT	"Genetic variation of HIV type 1 in four world Health Organization-			
RT	sponsored vaccine evaluation sites: generation of functional envelope			
RT	(glycoprotein 160) clones representative of sequence subtypes A, B, C,			
RL	and E. WHO Network for HIV Isolation and Characterization.";			
BL	AIDS Res. Hum. Retroviruses 10:1359-1368(1994).			
RN	(2)			

```

RP      SEQUENCE FROM N.A.
RC      STRAIN=014;
RX      MEDLINE=95194691; PubMed=7545977;
RA      WHO Global Programme on AIDS;
RT      "HIV type 1 variation in World Health Organization-sponsored vaccine
RT      evaluation sites: genetic screening, sequence analysis, and
RT      preliminary biological characterization of selected viral strains. WHO
RT      Network for HIV Isolation and Characterization."
RL      AIDS Res. Hum. Retroviruses 10:1327-1343(1994).
RN      [3]
RP      SEQUENCE FROM N.A.
RC      STRAIN=014;
RX      MEDLINE=96190564; PubMed=8627686;
RA      Gao F., Morrison S.G., Robertson D.L., Thornton C.L., Craig S.,
RA      Karlsson G., Sodroski J., Morgado M., Galvao-Castro B.,
RA      von Briesen H., Beddows S., Weber J., Sharp P.M., Shaw G.M.,
RA      Hahn B.H.;
RT      "Molecular cloning and analysis of functional envelope genes from
RT      human immunodeficiency virus type 1 sequence subtypes A through G. The
RT      WHO and NIAID Networks for HIV Isolation and Characterization."
RL      J. Virol. 70:1651-1657(1996).
RN      [4]
RP      SEQUENCE FROM N.A.
RC      STRAIN=014;
RA      Allen E.E.;
RL      Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
DR      EMBL; U08801; AAB05185.1;
DR      InterPro; IPR000328; Env_GP41.
DR      InterPro; IPR000777; GP120.
DR      Pfam; PF00516; GP120; 1.
DR      Pfam; PF00517; GP41; 1.
DR      AIDS; Coat protein; Glycoprotein; Polypotein; Transmembrane.
SQ      NON TER
FT      SEQUENCE 843 AA; 95679 MW; 14DB9A6D2F5FCE4 CRC64;

Query Match      87.8%; Score 2747; DB 15; Length 843;
Best Local Similarity 80.9%; Pred. No. 1,2e-226;
Matches 518; Conservative 23; Mismatches 37; Indels 62; Gaps 3

OY      2      EKLWTVYVYGVPMVEATTTTFCASDADAYDTEVNNWATHACVPTDPVQEVLENTE 61
DB      29      EKLWTVYVYGVPMVEATTTTFCASDADAYDTEVNNWATHACVPTDPVQEVLENTE 88
OY      62      HFNWKNMNVQEQMEDIIISLWDQSLKPCVKLTPLCA----- 98
DB      89      NFNWKNMNVQEQMEDIIISLWDQSLKPCVKLTPLCVLTINCTDLRATNTSSYMEGEIKN 148
OY      99      -----GDTSVITQACPXISFEPIPIHY 121
DB      149      CSFNITTSIKTKVVDYALFYKLDVVPIDNDNTSYRLINCNTSVITQACPXVFEPIPIHY 208
OY      122      CAPAGFALTKNDKTFNGKPGCKNVSTVQCTHGIRPVVSTOLLNGSLAEEVVRSDNF 181
DB      209      CTGPGFALQCNKKFNGTGCTNVSTVQCTHGIRPVVSTOLLNGSLAEEVVRSSNF 268
OY      182      TNNAKTIIVQKESVEINCTRPNNNTRKSIHIGPRAFTTGEIIGDIRQAHCNISRAKY 241
DB      269      TDNARVIIIVQNESVEINCTRPNNNTRKSIHIGPRAVYTTGQIIGDIRQAHCNISSTKN 328
OY      242      NDTLKQIYIKIREQENKTIYFNHSSGGDPETVWMSFNGCGEFPYCNSITQLFNSTNNNT 301
DB      329      NNTLRQITREKLRBQGNKTIYFNOSGGDPETVWMSFNGCGEFPYCNTTQLFNSTNNDTS 388
OY      302      EGSNNTBSNTTLPRIKQIINMGOVEKAWAPARIQRIORSSNITGLLFRDGSINEN 361
DB      389      TNNNTNGKTIITLPRIKQIIVNMGOVEKAWAPARIQRIORSSNITGLLFRDGSINEN 448
OY      362      -GTETFRGGGDMNDNRSELKYKVYVIEPLGVAPTCKRVRREKRAVG-IGAVFLG 419
DB      449      KTTETFRGGGDMNDNRSELKYKVYVIEPLGVAPTPKRRVVRREKRAVGTIGMFLG 508
OY      420      FLGAAGSTMGAASTLTYQARLLISGIYQOQNTLRATIDAQRMQLTYWGIKQIQARYL 479

```

```

Db      509 FLGTAGSTMGASITLTVQARQLSGIVQOQRNLLRAIEAQOHLJOLTVWGIKOLQARVL 568
Qy      480 AVERYLDGQOLLGIMGSGKLICTTAVPWNASWNSKSLDRIWNNMTWMEWEREIDNTSE 539
Db      569 AVERYLDGQOLLGIMGSGKLICTTAVPWNASWNSKSLDKIWNNTWMEWEREIDNTRE 628
Qy      540 IYTLIESQNOQEKNEOELLELDKXASIMNWFDTNNLWY 579
Db      629 IYTLIESQNOQEKNEELLELDKXASIMNWFDTNNLWY 668

```

Search completed: December 12, 2003, 12:32:39
 Job time : 32.7807 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 12:25:38 ; Search time 9.71792 Seconds
(without alignments)
2801.881 Million cell updates/sec

Title: US-10-032-162-15

Perfect score: 3129
Sequence: 1 VEKLVVTVYGVVWKEAT.....ELDKWASLWNPEDITWLMY 579

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues
1 number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2799.5	89.5	848	1 ENV_HV1JR	P20871 human immun
2	2764	88.3	847	1 ENV_HV1SI	P19550 human immun
3	2754.5	88.0	856	1 ENV_HV1SC	P05878 human immun
4	2724.5	87.1	861	1 ENV_HV1BR	P03377 human immun
5	2723	87.0	867	1 ENV_HV1J3	P12489 human immun
6	2719	86.9	856	1 ENV_HV1B1	P03375 human immun
7	2718.5	86.9	856	1 ENV_HV1W1	P31872 human immun
8	2715	86.8	856	1 ENV_HV1PV	P03376 human immun
9	2714	86.7	855	1 ENV_HV1A2	P03378 human immun
10	2706	86.5	843	1 ENV_HV1Y2	P35961 human immun
11	2702	86.4	856	1 ENV_HV1H2	P04578 human immun
12	2699	86.3	856	1 ENV_HV1H3	P04524 human immun
13	2696.5	86.2	851	1 ENV_HV1B8	P04582 human immun
14	2696	86.2	852	1 ENV_HV1BN	P12488 human immun
15	2693	86.1	856	1 ENV_HV1WM	O70626 human immun
16	2690	86.0	856	1 ENV_HV1LM	P05877 human immun
17	2672	85.4	847	1 ENV_HV1W2	P05880 human immun
18	2670	85.3	853	1 ENV_HV1W2	P19551 human immun
19	2663	85.1	865	1 ENV_HV1RH	P04579 human immun
20	2657.5	84.9	852	1 ENV_HV1S3	P19549 human immun
21	2633	84.1	855	1 ENV_HV1OY	P20888 human immun
22	2628.5	84.0	861	1 ENV_HV1KB	P18189 human immun
23	2625.5	83.9	868	1 ENV_HV1C4	P05879 human immun
24	2496	79.8	855	1 ENV_HV1Z6	P04580 human immun
25	2467	78.2	853	1 ENV_HV1Z2	P12487 human immun
26	2448	78.8	863	1 ENV_HV1Z8	P05882 human immun
27	2444.5	78.1	859	1 ENV_HV1MA	P04583 human immun
28	2443.5	78.1	856	1 ENV_HV1ZH	P05881 human immun
29	2428.5	77.6	846	1 ENV_HV1ND	P18799 human immun
30	2389	76.4	853	1 ENV_HV1EL	P04581 human immun
31	2060	65.8	854	1 ENV_HV1CZ	P17281 chimpanzee
32	1458.5	46.6	421	1 ENV_HV1N5	P12490 human immun
33	1427	45.6	460	1 ENV_HV1Z3	P12491 human immun

34	1108.5	35.4	854	1 ENV_S1VA1	Q02837 simian immu
35	1106	35.3	859	1 ENV_HV2ST	P20872 human immu
36	1096.5	35.0	851	1 ENV_HV2G1	P18040 human immu
37	1093	34.9	712	1 ENV_HV2S2	P32535 human immu
38	1087.5	34.8	768	1 ENV_S1VA1	P22757 simian immu
39	1077	34.4	877	1 ENV_S1VAG	P27977 simian immu
40	1076	34.4	858	1 ENV_HV2RO	P04577 human immu
41	1066.5	34.1	859	1 ENV_HV2CA	P24105 human immu
42	1065	34.0	856	1 ENV_HV2NZ	P05883 human immu
43	1060	33.9	865	1 ENV_S1VAT	P05886 simian immu
44	1058.5	33.8	851	1 ENV_HV2D1	P17755 human immu
45	1055	33.7	846	1 ENV_HV2SB	P12449 human immu

ALIGNMENTS

RESULT 1	ID	ENV_HV1JR	STANDARD;	PRT;	848 AA.
AC	P20871;				
DT	01-FEB-1991 (Rel. 17, Created)				
DT	01-FEB-1991 (Rel. 17, Last sequence update)				
DT	15-SEP-2003 (Rel. 42, Last annotation update)				
DE	Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].				
GN	ENV.				
OS	Human immunodeficiency virus type 1 (JRCSF isolate) (HIV-1).				
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.				
OX	NCBI_Taxid=11688;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Koyanagi S., Chen I.S.Y.;				
RL	Submitted (DEC-1988) to the HIV data bank.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by, and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; M38429; AAB03749.1; .				
DR	PDB; 1CB4; 1B-MAR-99				
DR	HIV; M38429; ENV\$JRCSF.				
DR	InterPro; IPR000328; Env GP41.				
DR	InterPro; IPR000777; GP120.				
DR	Pfam; PF00516; GP120; 1.				
DR	Pfam; PF00517; GP41; 1.				
KW	Aids; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal; 3D-structure.				
KW	SIGNAL				
FT	FT	3	503	32	
FT	CHAIN	504	848		EXTERIOR MEMBRANE GLYCOPROTEIN.
FT	DISULFID	53	73		TRANSMEMBRANE GLYCOPROTEIN.
FT	DISULFID	118	203		BY SIMILARITY.
FT	DISULFID	125	194		BY SIMILARITY.
FT	DISULFID	130	154		BY SIMILARITY.
FT	DISULFID	216	245		BY SIMILARITY.
FT	DISULFID	226	237		BY SIMILARITY.
FT	DISULFID	294	328		BY SIMILARITY.
FT	DISULFID	374	437		BY SIMILARITY.
FT	DISULFID	381	410		BY SIMILARITY.
FT	CARBOHYD	87	87		N-LINKED (GLCNAC. . .)
FT	CARBOHYD	114	134		N-LINKED (GLCNAC. . .)
FT	CARBOHYD	137	137		N-LINKED (GLCNAC. . .)
FT	CARBOHYD	153	153		N-LINKED (GLCNAC. . .)
FT	CARBOHYD	157	157		N-LINKED (GLCNAC. . .)
FT	CARBOHYD	165	157		N-LINKED (GLCNAC. . .)
FT	CARBOHYD	185	185		N-LINKED (GLCNAC. . .)
FT	CARBOHYD	195	195		N-LINKED (GLCNAC. . .)
FT	CARBOHYD	228	228		N-LINKED (GLCNAC. . .)
FT	CARBOHYD	239	239		N-LINKED (GLCNAC. . .)

```

FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 274 274 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 382 382 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 388 388 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 403 403 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 440 440 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 453 453 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 603 603 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 608 608 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 617 617 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 629 629 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 808 808 N-LINKED (GLCNAC. . .) (POTENTIAL)
SEQUENCE 848 AA; 96475 MW; 2076751227EC3F3 CRC64;

Query Match 89.5%; Score 2799.5; DB 1; Length 848;
Best Local Similarity 82.2%; Pred. No. 7.1e-219;
Matches 530; Conservative 20; Mismatches 28; Indels 67; Gaps 3;

Qy 1 VEKLMVTVYVGVVPWKEATTTTLCASDAKAYDTEVHNWATHACVPTDPNPOEVLENTV 60
Db 30 VEKLMVTVYVGVVPWKEATTTTLCASDAKADTEVHNWATHACVPTDPNPOEVLENTV 89
Qy 61 EHFNMKNMWEQOEEDIIISLMDSLKPCVKLTPLCGA----- 98
Db 90 EDFNMKNMWEQOEEDVINLMDSLKPCVKLTPLCVTLNCKDVNATNTSSSEGMERG 149
Qy 99 -----GQDSVTVQACPKISFE 115
Db 150 EIKNCSFNITYSIRDKVQKEYALFYKLDVPIDNKNTRYRLISGNTSVITQACPKSFE 209
Qy 116 PIPHYCAPAGFALIKCNDKTFNGKPCKNVSTVOCTGIRPVVSTOLLNGLSLAEERV 175
Db 210 PIPHYCAPAGFALIKCNDKTFNGKPCKNVSTVOCTGIRPVVSTOLLNGLSLAEERV 269
Qy 176 IRSDFNNAKTIIVOLKESVEINCTRPNNTRKSIHIGPRAFYTTGEIIGDIRQAHN 235
Db 270 IRSDFNNAKTIIVOLKESVYKINCTRPNNTRKSIHIGPRAFYTTGEIIGDIRQAHN 329
Qy 236 ISRAKNDTKQIYKLEQEBNKTIIVNHSNGGPEIIVMSFNGGFFFCNSLTQFNLS 295
Db 330 ISRAQNMNTLKOIVEKLEQEPFNKTIIVTSSGGDPEIIVMSFNGGFFFCNSLTQFNLS 389
Qy 296 TWNNNTSGSNNTGEG-TITLPCRIKOIIMMOEVCAYAPPIRQOICSSNITGLLTLR 354
Db 390 TW-NDTESSSGTSGNDTIIIPCRIOIIMMOEVCAYAPPIRQOICSSNITGLLTLR 448
Qy 355 DQGINENGEIIFRPGGDMRDWRSELYKYVVKIIEPLGVAPTKCRREVVOREKAVIG 414
Db 449 DQGNESIEIIFRPGGDMRDWRSELYKYVVKIIEPLGVAPTKCRREVVOREKAVIG 508
Qy 415 AVFLGFLGAAGSTWAAAMTTLTQARLLLSGIVOONNLRAIEQOQMLQUTWIGITQL 474
Db 509 ALFLGFLGAAGSTWAAAMTTLTQARLLLSGIVOONNLRAIEQOQMLQUTWIGITQL 568
Qy 475 QARVLAVERLYGDOOLLGIWGGSLICTTAPVPMASNSKSLDIRIMNMNTMEWEREID 534
Db 569 QARVLAVERLYKDOOLMGKGGSLICTTAPVPMNTSSNSKSLDIRIMNMNTMEWEREID 628
Qy 535 NYTSEIYTLIESQOQEKNEBELLELDKMASLMMWFDTLWMLWY 579
Db 629 NYTNTIYTLIESQOQEKNEBELLELDKMASLMMWFDTLWMLWY 673

```

```

AC P19550;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [contains: Exterior membrane
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (SF162 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11691;
RX MEDLINE=90347835; PubMed=2384920;
RA Cheng-Mayer C., Quiroga M., Tung J.W., Levy J.;
RT "Viral determinants of human immunodeficiency virus type 1 T-cell or
RL J. Virol. 64:4390-4398(1990).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M65024; AAA5072.1; -.
CC PDB: 10B5; 15-MAY-97.
CC HIV; M38428; ENV$SF162.
CC InterPro; IPR000328; Env_GP41.
CC InterPro; IPR000777; GP120.
CC Pfam; PF00516; GP120; 1.
CC DR Pfam; PF00517; GP41; 1.
CC KW AIDS; Coat protein; Glycoprotein; Glycoprotein; Transmembrane; Signal;
CC 3D-structure.
CC SIGNAL 1 29
CC CHAIN 30 502
CC FT DIULFID 53 73
CC FT DIULFID 118 203
CC FT DIULFID 125 194
CC FT DIULFID 130 155
CC FT DIULFID 216 245
CC FT DIULFID 226 237
CC FT DIULFID 294 328
CC FT DIULFID 374 435
CC FT DIULFID 381 408
CC FT CARBOHYD 87 87
CC FT CARBOHYD 135 135
CC FT CARBOHYD 154 154
CC FT CARBOHYD 186 186
CC FT CARBOHYD 195 195
CC FT CARBOHYD 232 232
CC FT CARBOHYD 239 239
CC FT CARBOHYD 260 260
CC FT CARBOHYD 274 274
CC FT CARBOHYD 293 293
CC FT CARBOHYD 299 299
CC FT CARBOHYD 329 329
CC FT CARBOHYD 336 336
CC FT CARBOHYD 352 352
CC FT CARBOHYD 382 382
CC FT CARBOHYD 388 388
CC FT CARBOHYD 392 392
CC FT CARBOHYD 398 398
CC FT CARBOHYD 401 401
CC FT CARBOHYD 438 438
CC FT CARBOHYD 454 454
CC FT CARBOHYD 602 602
CC FT CARBOHYD 607 607
CC FT CARBOHYD 616 616
CC FT CARBOHYD 628 628
CC SEQUENCE 847 AA; 96135 MW; 0A901317FD7FF2AB CRC64;

```


Query Match 88.3%; Score 2764; DB 1; Length 847;
 Best Local Similarity 81.7%; Pred. No. 5.3e-216;
 Matches 527; Conservative 16; Mismatches 34; Indels 68; Gaps 4;

1 VEKLVTVYVGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENT 60
 30 VEKLVTVYVGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENT 89
 61 EHFNNMKNMVEQOMEDIISLWDQSLKPCVKLTPLCGA----- 98
 90 EHFNNMKNMVEQOMEDIISLWDQSLKPCVKLTPLCVTLHCTNLKNATNKSNNKEMDR 149
 99 -----GCPTSVITQACPKISPE 115
 150 GEIKNCSFKVTSIRNNQKEVALFYKLDVPIIDNDNTSYKLNCTNITSYITQACPKVSPE 209
 116 PIPHYCAPAGFAILLKNDKTNKGKPCKNVSTVOCTHGIRPVSTOLLNGLAEEVY 175
 210 PIPHYCAPAGFAILLKNDKTNKGKPCKNVSTVOCTHGIRPVSTOLLNGLAEEVY 269
 176 IRSDFNNAKTIIVOLKESVEINCTRPNNNRKSHIGRGAFTTGGIIGDIOAHN 235
 270 IRSDFNNAKTIIVOLKESVEINCTRPNNNRKSHIGRGAFTTGGIIGDIOAHN 329
 236 ISRAKMDTLKQIVIKLREOFENKTIIVFNHSGGPELVMHSGGPEFYCNSTOLEFS 295
 330 ISGEKMNNTLKQIVIKLQAFQNKTIIVFQSSGGPELVMHSGGPEFYCNSTOLEFS 389
 296 TWNNNTBEGSNTEGTTTLCRIKQIINNMOEYKAMTAPIRGOIRCSNITGILLTRD 355
 390 TWNNNTBEGSNTEGTTTLCRIKQIINNMOEYKAMTAPIRGOIRCSNITGILLTRD 447
 356 GGIN-ENGELPRGGGDMRWSELYKYKVKLEPGVATKCKRPVQEKRAVIG 414
 448 GCKEISNTTELPFGGDMRWSELYKYKVKLEPGVATKCKRPVQEKRAVIG 507
 415 AVFLGFLGAGSTWGAASMTLTVQARLLSGIVQOONNLRAIEAQRMQLTWGICKOL 474
 508 AMFLGFLGAGSTWGAASMTLTVQARLLSGIVQOONNLRAIEAQRMQLTWGICKOL 567
 475 QARVLAVERRYLQDQQLLGWCGSGKLICTTAVPNNASMSKSLDRIINNMTMEMEREID 534
 568 QARVLAVERRYLQDQQLLGWCGSGKLICTTAVPNNASMSKSLDRIINNMTMEMEREID 627
 535 NYSIYTLIESQNOQEKNEBELLEDKMSLMMWPTTMMWY 579
 628 NYSIYTLIESQNOQEKNEBELLEDKMSLMMWPTTMMWY 672

RESULT 3
 ENV_HVISC STANDARD; PRT; 856 AA.
 AC P05878;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
 GN HIV.
 OS Human immunodeficiency virus type 1 (SC isolate) (HIV-1).
 OC Viruses; Retrovirus; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=88219542; PubMed=3369091; Aldovini A., Collalti E., Farrel K., Wong-Staal F., Gallo R.C., Reitz M.S. Jr.;
 RT "Envelope sequences of two new United States HIV-1 isolates.";
 RL Virology 164:531-536(1988).
 CC -1- MISCELLANEOUS: THE SC ISOLATE WAS TAKEN FROM AN ARC PATIENT IN 1984 IN SOUTHERN CALIFORNIA.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, M17450; -; NOT_ANNOTATED_CDS.
 DR HIV, M17450; ENVSSC.
 DR InterPro; IPR000328; Env_GP41.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 KM AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal.
 KW SIGNAL.
 FT SIGNAL 1 29
 FT CHAIN 30 510
 FT CHAIN 511 856
 FT SITE 760 760
 FT DISULFID 53 73
 FT DISULFID 118 206
 FT DISULFID 125 197
 FT DISULFID 130 160
 FT DISULFID 219 247
 FT DISULFID 228 239
 FT DISULFID 296 330
 FT DISULFID 376 439
 FT DISULFID 383 412
 FT CARBOHYD 87 87
 FT CARBOHYD 129 129
 FT CARBOHYD 135 135
 FT CARBOHYD 140 140
 FT CARBOHYD 143 143
 FT CARBOHYD 159 159
 FT CARBOHYD 163 163
 FT CARBOHYD 189 189
 FT CARBOHYD 198 198
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 302 302
 FT CARBOHYD 331 331
 FT CARBOHYD 338 338
 FT CARBOHYD 354 354
 FT CARBOHYD 360 360
 FT CARBOHYD 384 384
 FT CARBOHYD 394 394
 FT CARBOHYD 400 400
 FT CARBOHYD 405 405
 FT CARBOHYD 442 442
 FT CARBOHYD 457 457
 FT CARBOHYD 611 611
 FT CARBOHYD 616 616
 FT CARBOHYD 625 625
 FT CARBOHYD 637 637
 FT CARBOHYD 674 674
 FT CARBOHYD 816 816
 SQ SEQUENCE 856 AA; 97055 MW; DAFDA600EBA7A08 CRC64;

Query Match 88.0%; Score 2754.5; DB 1; Length 856;
 Best Local Similarity 80.1%; Pred. No. 3.2e-215;
 Matches 523; Conservative 26; Mismatches 27; Indels 77; Gaps 6;

2 EKLWTVYVGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENT 61
 31 EKLWTVYVGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENT 90
 62 EHFNNMKNMVEQOMEDIISLWDQSLKPCVKLTPLCGA----- 98


```

Db 91 NFNMMKNMNVQWQHEHDIISLMDQSLKPCVKLTPLCVLTNLTNLTNDSTNATNTSSNRG 150
Qy 99 -----GCDTSVITQACPKIS 113
Db 151 KMEGCEMNCSPNITTSIRSKVQKEYALFYKLDVVPIDNTSYTLTNCSTSYITQACPKIS 210
Qy 114 FEPPIPIHYCAPAGFAILKNDKTFNGKPCCKNVSTVOCTHGIRPVYSTOLLNGLSGLAEE 173
Db 211 FEPPIPIHYCA-RWFAILNCKNKKEFGTGPCTNVSTVOCTHGIRPVYSTOLLNGLSGLAEE 269
Qy 174 VVIRSDNFTNNAKTIIVOLKESVEINCTRNPNNTKRSIHIGPRAFYATGDIIDIRQAH 233
Db 270 VVLRSENTDPAKTIIVOLKESVEINCTRNPNNTKRSIHIGPRAFYATGDIIDIRQAH 329
Qy 234 CNIRAKKNDTLKQIVILKREOPENKTIIVFNHSSGQDEIIVHSPNCGSEFFYCNSTOLF 293
Db 330 CNIRAKKNDTLKQIVILKREOPENKTIIVFNHSSGQDEIIVHSPNCGSEFFYCNSTOLF 389
Qy 294 NSTWNNNTSGSNTEGN-TITLPCRKQIIMMOWEVRWYAPPIRQIRCSNITGILL 352
Db 390 SSTW-NGTESSNNTGNDTITLPCRKKEIIMMOWEVRWYAPPIRQIRCSNITGILL 448
Qy 353 TRDGGINENG-----TEIFRPGGDMRDNRSELYKYVVKIEPLGVAPTYAKRRVQRE 407
Db 449 TRDGGNSKNGSKNTEIFRPGGDMRDNRSELYKYVVKIEPLGVAPTYAKRRVQRE 508
Qy 408 KRAVG-IGAVPLGLFGAAGSTMGASMTLTVQARLLSGIVQOQNNLRAIEAOQMLQ 466
Db 509 KRAVGITGAMPLGLFGAAGSTMGATSMTLTVQARLLSGIVQOQNNLRAIEAOQMLQ 568
Qy 467 TWVGIKQOLAVLAVERYLGDQQLGIGWCSGKLICTTAVPMNASWSNKSIDRIMNNMTW 526
Db 569 TWVGIKQOLAVLAVERYLROQLLIGIGCSGKLICTTAVPMNASWSNKSIDRIMNNMTW 628
Qy 527 MEMREIDNTSYSELYTILIESQOQEKNOEOLBELDKRASIIMNFDTINWIMY 579
Db 629 MEMREIDNTSYLYTILIESQOQEKNOEOLBELDKRASIIMNFDTINWIMY 681

RESULT 4
ENV_HV1BR STANDARD; PRT; 861 AA.
ID ENV_HV1BR
AC P03377;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
ENY.
CC Human immunodeficiency virus type 1 (BRU isolate) (HIV-1).
CC Viruses; Retrovirus; Retroviridae; Lentivirus.
CC NCBI_TaxID=11686;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8509333; PubMed=2981635;
RA Main-Hobson S., Sonigo P., Dano O., Cole S., Allison M.;
RT "Nucleotide sequence of the AIDS virus, LAV.";
RL Cell 40:9-17(1985).

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; K02013; AAB59751.1; -
DR EMBL; A04321; CA00352.1; -
DR PIR; A03975; VCLLV.
DR PDB; 1ERF; 20-FEB-02.
DR HIV; K02013; ENV5BRU.
DR InterPro; IPR000328; Env_GP41.

```

```

DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal;
KW 3D-structure.
FT SIGNAL 1
FT CHAIN 31
FT CHAIN 516
FT DISULFID 517
FT DISULFID 54
FT DISULFID 74
FT DISULFID 119
FT DISULFID 126
FT DISULFID 131
FT DISULFID 131
FT DISULFID 162
FT DISULFID 223
FT DISULFID 233
FT DISULFID 301
FT DISULFID 383
FT DISULFID 390
FT CARBOHYD 88
FT CARBOHYD 136
FT CARBOHYD 141
FT CARBOHYD 146
FT CARBOHYD 161
FT CARBOHYD 165
FT CARBOHYD 191
FT CARBOHYD 191
FT CARBOHYD 202
FT CARBOHYD 235
FT CARBOHYD 239
FT CARBOHYD 246
FT CARBOHYD 267
FT CARBOHYD 281
FT CARBOHYD 294
FT CARBOHYD 300
FT CARBOHYD 306
FT CARBOHYD 337
FT CARBOHYD 344
FT CARBOHYD 361
FT CARBOHYD 391
FT CARBOHYD 397
FT CARBOHYD 402
FT CARBOHYD 411
FT CARBOHYD 453
FT CARBOHYD 453
FT CARBOHYD 468
FT CARBOHYD 616
FT CARBOHYD 621
FT CARBOHYD 630
FT CARBOHYD 642
FT CARBOHYD 679
FT CARBOHYD 755
FT CARBOHYD 821
SQ SEQUENCE 861 AA; 97487 MW; 04DE2B4D4E4FD63A CRC64;

Query Match 87.1%; Score 2724.5; DB 1; Length 861;
Best Local Similarity 79.6%; Pred. No. 8.6e-213;
Matches 522; Conservative 25; Mismatches 30; Indels 79; Gaps 6;

Qy 2 EKLWVTVYGVVVKRATTTLFCASDAKAYDTEVHNWMAHFAVCPTDPNPOEVLNTE 61
Db 32 EKLWVTVYGVVVKRATTTLFCASDAKAYDTEVHNWMAHFAVCPTDPNPOEVLNTE 91
Qy 62 HFNMMKNMNVQWQHEHDIISLMDQSLKPCVKLTPLCGA----- 98
Db 92 NFNMMKNMNVQWQHEHDIISLMDQSLKPCVKLTPLCGA----- 151
Qy 99 -----GCDTSVITQACPKIS 110
Db 152 MMEKEIKNCSFNISTIRSKVQKEYAFYKLDIIPIDNTSYTLTNCSTSYITQACPKIS 211
Qy 111 KISFEPPIHYCAPAGFAILKNDKTFNGKPCCKNVSTVOCTHGIRPVYSTOLLNGLSGLA 170
Db 212 KISFEPPIHYCAPAGFAILKNDKTFNGKPCCKNVSTVOCTHGIRPVYSTOLLNGLSGLA 271
Qy 171 EEVVIRSDNFTNNAKTIIVOLKESVEINCTRNPNNTKRSIHIGPRAFYATGDIIDIRQAH 228

```

Db	272	EEBVRISANFTDNAKTTIIVQLNQSVETICTRPNNNTRKSIRIQRPBGAFTYIGK-IGN	330
Qy	229	IRQACNISBRAMNDTKQIVYKARQF-EAKTIYVNHSSGGDEPIVMHSFNGCGEPFYC	287
Db	331	IRQACNISBRAMNATLXQIAKLTNEQFQGNKTTIIFKQSSGGPEIVTSHFNGCGEPFYC	390
Qy	288	NSTOLFNSFTWNN--NTEGSSNTEG-NTTLPCRIOIIMNQEYKAYAPPIROIQC	343
Db	391	NSTOLFNSFTWNSFTWSTBGSNNTEGDTTLPCRIOFINNQEYKAYAPPIROIQC	450
Qy	344	SSNTGLLTPRDGINENGTEIFRPGCGMDRDNMSSELYKYKVKTEPIGVAPTKCRRY	403
Db	451	SSNTGLLTPRGGNNNGNSEIFRPGCGMDRDNMSSELYKYKVKTEPIGVAPTKCRRY	510
Qy	404	VQREKRAVGICGLFPGFGAAGSTGKASMTLLTVQARLLSGIVQOQNNLLRAIEAQRM	463
Db	511	VQREKRAVGICGLFPGFGAAGSTGKASMTLLTVQARLLSGIVQOQNNLLRAIEAQRM	570
Qy	464	LQLTWVGKQLOARLAVERYLQDQQLIGWCGSGULICTTAVPMNASVSNLSLDTIMNN	523
Db	571	LQLTWVGKQLOARLAVERYLQDQQLIGWCGSGULICTTAVPMNASVSNLSLDTIMNN	630
Qy	524	MTWMEBERIDNYTSEIYVTLIEESONQOEKNEQLELDDKMASLMMPTITWLMY	579
Db	631	MTWMEBERIDNYTSEIYVTLIEESONQOEKNEQLELDDKMASLMMPTITWLMY	686
ENV_HV1J3			
ENV_HV1J3			
ID	ENV_HV1J3	STANDARD;	PRT; 867 AA.
AC	P12489;		
DT	01-OCT-1989	(Rel. 12, Created)	
DT	01-OCT-1989	(Rel. 12, Last sequence update)	
DT	15-JUL-1999	(Rel. 38, Last annotation update)	
DE	Envelope glycoprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].		
GN	ENV.		
OS	Human immunodeficiency virus type 1 (H3 isolate).		
OC	Virusess; Retroid virusess; Retroviridae; Lentivirus.		
OX	NCBI_TaxID=11694;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=89352108; PubMed=2669897;		
RA	Komiyama N., Hattori N., Inoue J., Sakuma S., Kurimura T., Yoshida M.,		
RT	"Nucleotide sequences of gag and env genes of a Japanese isolate of		
RT	HIV-1 and their expression in bacteria."		
RL	AIDS Res. Hum. Retrovirus 5:411-419(1989).		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation between the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; M21138; AAB03526.1; -.		
DR	HIV; M21138; ENVSJ3.		
DR	InterPro; IPR000328; Env GP41.		
DR	InterPro; IPR000777; GP120.		
DR	Pfam; PF00516; GP120.1.		
DR	Pfam; PF00517; GP41.1.		
KW	AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal.		
FT	SIGNAL		
FT	CHAIN	1	30
FT	CHAIN	517	516
FT	DISULFID	53	73
FT	DISULFID	118	217
FT	DISULFID	125	208
FT	DISULFID	130	160
FT	DISULFID	230	259
FT	DISULFID	240	251
FT	CHAIN	31	516
FT	CHAIN	517	867
FT	DISULFID	53	73
FT	DISULFID	118	217
FT	DISULFID	125	208
FT	DISULFID	130	160
FT	DISULFID	230	259
FT	DISULFID	240	251

FT	DISLFTID	308	342	BY SIMILARITY.
FT <td>DISLFTID</td> <td>388</td> <td>457</td> <td>BY SIMILARITY.</td>	DISLFTID	388	457	BY SIMILARITY.
FT <td>DISLFTID</td> <td>395</td> <td>430</td> <td>BY SIMILARITY.</td>	DISLFTID	395	430	BY SIMILARITY.
FT <td>CARBOHYD</td> <td>87</td> <td>87</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL).</td>	CARBOHYD	87	87	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>135</td> <td>135</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL).</td>	CARBOHYD	135	135	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>140</td> <td>140</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL).</td>	CARBOHYD	140	140	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>143</td> <td>143</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL).</td>	CARBOHYD	143	143	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>159</td> <td>159</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL).</td>	CARBOHYD	159	159	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>163</td> <td>163</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL).</td>	CARBOHYD	163	163	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>188</td> <td>188</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL).</td>	CARBOHYD	188	188	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>189</td> <td>189</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL).</td>	CARBOHYD	189	189	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>199</td> <td>199</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL).</td>	CARBOHYD	199	199	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>209</td> <td>209</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL).</td>	CARBOHYD	209	209	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>246</td> <td>246</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL).</td>	CARBOHYD	246	246	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>253</td> <td>253</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL).</td>	CARBOHYD	253	253	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>274</td> <td>274</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL).</td>	CARBOHYD	274	274	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>288</td> <td>288</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL).</td>	CARBOHYD	288	288	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>307</td> <td>307</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL).</td>	CARBOHYD	307	307	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>350</td> <td>350</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL).</td>	CARBOHYD	350	350	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>366</td> <td>366</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL).</td>	CARBOHYD	366	366	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>372</td> <td>372</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL).</td>	CARBOHYD	372	372	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>396</td> <td>396</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL).</td>	CARBOHYD	396	396	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>402</td> <td>402</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL).</td>	CARBOHYD	402	402	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>408</td> <td>408</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL).</td>	CARBOHYD	408	408	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>412</td> <td>412</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL).</td>	CARBOHYD	412	412	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>418</td> <td>418</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL).</td>	CARBOHYD	418	418	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>423</td> <td>423</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL).</td>	CARBOHYD	423	423	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>460</td> <td>460</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL).</td>	CARBOHYD	460	460	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>475</td> <td>475</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL).</td>	CARBOHYD	475	475	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>622</td> <td>622</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL).</td>	CARBOHYD	622	622	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>627</td> <td>627</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL).</td>	CARBOHYD	627	627	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>636</td> <td>636</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL).</td>	CARBOHYD	636	636	N-LINKED (GLCNAC . . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>648</td> <td>648</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL).</td>	CARBOHYD	648	648	N-LINKED (GLCNAC . . .) (POTENTIAL).
SEQ	SEQUENCE	867 AA;	98399 MM;	5F2310146B8B8680 CRC64;
Query Match	Best Local Similarity	87.0%;	Score 2123;	DB 1; Length 867;
Matches 521;	Conservative	20;	Mismatches 36;	Indels 86; Gaps 4
OY	2	EKLWTVYVYGVVPMKEATTTLCASDAAYOTEVNWNATHACVPTDPPPOSVLENTE	61	
DB	31	EOLWTVYVYGVVPMKEAATTTLCASDAAYVDEVNWNATHACVPTDPPPOSVLENTE	90	
OY	62	HFNNMKNMNVOMOBEDILSLMDOSLKPCVKLTPLCGA-----	98	
DB	91	KFNMMKNMNVOMOBEDILSLMDOSLKPCVKLTPLCGA-----	150	
OY	99	-----GCDT	102	
DB	151	KMEKGEMKNCSFNITTSIRDVKQEHALFYKHDVVPINNSTKDNINKDNSTRYRLISCVT	210	
OY	103	SVITQACPKISFEPIPIHYCAPAGPALIKNDKPNFGNGPCNVSTVQCTHRIKPVSTO	162	
DB	211	SVITQACPKISFEPIPIHYCAPAGPALIKNDKPNFGNGPCNVSTVQCTHRIKPVSTO	270	
OY	163	LLNLGSLAEEBVRVRSDFNFTNNAKTIYIQLKESVEINCTRPNNNRKSHIGPGAFYTT	222	
DB	271	LLNLGSLAEEBVRVRSDFNFTNNAKTIYIQLKESVEINCTRPNNNRKSHIGPGAFYTT	330	
OY	223	GEIIGDIRQAHCSISRAKNDTLKOIVIKLBEPENKTIIVFNHSSGGDEIIVMHSFNCGG	282	
DB	331	KQIAGDLRQAHCSINIRARMAWTAKQIVGLKARKQFVNKTIIVFNHSSGGDEIIVMHSFNCGG	390	
OY	283	EFFCNSTQULFNSTWN-----NNTBGSNNTBGN-TTTPCRLKOIINNMQEPKAMVAP	336	
DB	391	EFFCNSTQULFNSTWN-----NNTBGSNNTBGN-TTTPCRLKOIINNMQEPKAMVAP	450	
OY	337	IRGOIRGCSNITGTLITRDGGINENGTEIFPPGGGDMDNMRSELKYKVVYIEPLGVAP	396	
DB	451	IRGOIRGCSNITGTLITRDGGINENGTEIFPPGGGDMDNMRSELKYKVVYIEPLGVAP	510	
OY	397	TKCRRVVRQBRKAVGIGAVELGELGAAGSTMGASMTLVQARLLLSGIVQOONNLLRA	456	

Db 511 TKARRVVRREKRAVIGAVELGFGAAGTGWG-ASMTLTVQARLLLSGIVQOONNLRA 569
 Qy 457 IEAQQRMQLTVMGKIKOLQARVLAVERLGDQQLIGWGSGKLIICCAVPAPNWSKKS 516
 Db 570 IEGQOHLQTLTVMGKIKOLQARVLAVERLGDQQLIGWGSGKLIICCAVPAPNWSKKS 629
 Qy 517 LDRIMNNMTWMEWESEIDNYSYITLTLEESONQOEKNEQELBLDKNASLMMNPDITNW 576
 Db 630 LEEIMDNMTWMEWESEIDNYSYITLTLEESONQOEKNEQELBLDKNASLMMNPDITNW 689
 Qy 577 LMY 579
 Db 690 LMY 692

RESULT 6
 ID ENV_HV1B1 STANDARD, PRT, 856 AA.
 AC P03375;
 21-JUL-1986 (Rel. 01, Created)
 21-JUL-1986 (Rel. 01, Last sequence update)
 15-JUL-1999 (Rel. 38, Last annotation update)
 Envelope polypeptide GP160 precursor (Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)).
 DE Envelope polypeptide GP160 precursor (Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)).
 GN ENV.
 OS Human immunodeficiency virus type 1 (BH10 isolate) (HIV-1).
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11678;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=8511123; PubMed=2578615;
 RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R., Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A., Baumesister K., Ivanoff L., Pettey S.R., Jr., Pearson M.L., Latcheberger J.A., Papas T.S., Ghrayeb J., Chang N.T., Gallo R.C., Wong-Staal F.;
 RA "Complete nucleotide sequence of the AIDS virus, HTLV-III.";
 RT Nature 313:277-284(1985).
 RL [2]
 RP DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.
 RX MEDLINE=90285159; PubMed=2355006;
 RA Leonard C.K., Spellman M.W., Riddle L., Harris R.J., Thomas J.N., Gregory T.J.;
 RA "Assignment of intrachain disulfide bonds and characterization of potential glycosylation sites of the type 1 recombinant human immunodeficiency virus envelope glycoprotein (gp120) expressed in Chinese hamster ovary cells.";
 RT J. Biol. Chem. 265:10373-10382(1990).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL: M15654; AAA44205.1;
 DR PIR: A03973; VCLH3.
 DR HIV: M15654; ENVBHI02.
 DR InterPro: IPR000328; Env_GP41.
 DR InterPro: IPR000777; GP120.
 DR Pfam: PF00516; GP120.1.
 DR Pfam: PF00517; GP41.1.
 DR AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal.
 KW Signal.
 FT CHAIN 1 30
 FT CHAIN 511 511 EXTERIOR MEMBRANE GLYCOPROTEIN.
 FT CHAIN 856 856 TRANSMEMBRANE GLYCOPROTEIN.
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157

FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT CARBOHYD 356 356
 FT CARBOHYD 386 386
 FT CARBOHYD 392 392
 FT CARBOHYD 397 397
 FT CARBOHYD 406 406
 FT CARBOHYD 448 448
 FT CARBOHYD 463 463
 FT CARBOHYD 611 611
 FT CARBOHYD 616 616
 FT CARBOHYD 625 625
 FT CARBOHYD 637 637
 FT CARBOHYD 674 674
 FT CARBOHYD 750 750
 FT CARBOHYD 816 816
 SQ SEQUENCE 856 AA; 97224 MW; 0BFBI1A18931B27 CRC64;

Query Match 86.9%; Score 2719; DB 1; Length 856;
 Best Local Similarity 80.0%; Pred. No. 2,4e-212;
 Matches 521; Conservative 28; Mismatches 28; Indels 74; Gaps 6;

Qy 2 EKLWVTVYGVVWKEATTLTFCASDAKAYDTEVHNWATHACVPTDPNQEVLNVTE 61
 Db 32 EKLWVTVYGVVWKEATTLTFCASDAKAYDTEVHNWATHACVPTDPNQEVLNVTE 91
 Qy 62 HFNWKKNNMVEQNEEDIIISLMDQSLKCVKLTPLCGA----- 98
 Db 92 NFNWKKNNMVEQNEEDIIISLMDQSLKCVKLTPLCVSLKCTDLKNDNTNSSSGRMIMK 151
 Qy 99 -----GCDTSVITQACPKISFE 115
 Db 152 GEINCSFNISTSRGVQKEAFYKLDIIPINDTTSYTLNSCNTSVITQACPKISFE 211
 Qy 116 PIPHYCAPAGFALIKCNKTPNGKGPCKNVSTVQCTHGIRPVVSTOLLNLSLAEEBV 175
 Db 212 PIPHYCAPAGFALIKCNKTPNGKGPCKNVSTVQCTHGIRPVVSTOLLNLSLAEEBV 271
 Qy 176 IRSNPFNNAKTIIIVOLKESVEINCTRPNNTRKSHI--GGPAAFTTBEIIGDIOAH 233
 Db 272 IRSNPFNNAKTIIIVOLKESVEINCTRPNNTRKSHI--GGPAAFTTBEIIGDIOAH 330
 Qy 234 CNISRAKNDPLTKOIVLKEBOF-ENKTIIVFNHSSGSDPEIVMHSFNGGGEFFYCNSTOL 292
 Db 331 CNISRAKNDPLTKOIVLKEBOF-ENKTIIVFNHSSGSDPEIVMHSFNGGGEFFYCNSTOL 390
 Qy 293 FNSWTNN---NTEGSNNTEG-NTITLPCRIKOIINMVEGKAMYPADPIRGQIRCSSNIT 348
 Db 391 FNSWTNN---NTEGSNNTEG-NTITLPCRIKOIINMVEGKAMYPADPIRGQIRCSSNIT 450
 Qy 349 GLILTRDGGINENGTETFRPGGDMRDNMSELKYKYVYKIEPVGAVPTCKRRVVRREK 408
 Db 451 GLILTRDGGINENGTETFRPGGDMRDNMSELKYKYVYKIEPVGAVPTCKRRVVRREK 510

FT	CARBOHYD	331	331	N-LINKED (GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	338	338	N-LINKED (GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	354	354	N-LINKED (GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	360	360	N-LINKED (GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	390	390	N-LINKED (GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	394	394	N-LINKED (GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	404	404	N-LINKED (GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	447	447	N-LINKED (GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	459	459	N-LINKED (GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	611	611	N-LINKED (GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	616	616	N-LINKED (GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	625	625	N-LINKED (GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	637	637	N-LINKED (GLCNAC . .)	(POTENTIAL)
SQ	SEQUENCE	856 AA;	97526 MM;	DB68D1E4C04D83 CRC64;	

Query Match 86.9%; Score 2718.5; DB 1; Length 856;
Best Local Similarity 79.9%; Pred. No. 2.6e-212;
Matches 522; Conservative 24; Mismatches 32; Indels 75; Gaps 7

Oy	1	VKKLVMTYYGVPPWKEKNTTLTFCASDAKAYDEVTHVNWMAHACYPTDPNPOEVLENTV	60
Dd	30	VEQLMVTYVVYGVPPWKKAETITLTFCASDKAKASTEHKRWATHACVPINPNPOEVLENTV	89
Oy	61	EHPFMKKNNWEQKHQIEDISLMDSLKPCVLTPLC-----GAG--	99
Dd	90	ENFPMKKNNWEQHHEIDISLMDSLKPCVXLTPLCVTLNCITDKNIIDMENKTIIIGSEBV	149
Oy	100	-----CDTSVITQAQPKRS	113
Dd	150	KNGCFNITITSIRDKVKHEVALFYKLADVPIKSNDSSITYRRLIHCHTSVITQAQSKYS	209
Oy	114	FEPPIPIHCAPAGAAILKCNDKTNGCKPCXONSVTOCTHGIRPVSTOLLNSGLAEIE	173
Dd	210	FEPPIPIHCAPAGAAILKCNDKCKFNKGTCPTNVSTVOCTHGIRPVSTOLLNSGLAEIE	269
Oy	174	IVVISDNFTNAKTIIVOLKESVEINCRPNPNN--RKSIIHGPGRAFTTEIIGDIRQA	232
Dd	270	IIVISENFTDNAKTIIVHLNESVEINCRPNNNRRRIHIHGPGRAF--TESIRGINIROA	328
Oy	233	HGNISRKAMNDTLKOQIVIKLEEPEKNKIIVNHSSGGDEPEIVMHSFNGCGEFYCNSTOL	292
Dd	329	HGNISRKAMNNTLKOQIVEKLEEGPKNTIIVNHSSGGDEPEIVMHSFNGCGEFYCDSTOL	388
Oy	293	FNSFWN--NNTSGSNNT--GNTITLPCKIKOIINMWQEVGKAMYAPPINGOIRCSSNI	347
Dd	389	FNSFWNWTGISTEGANNTEEENGDTITLPCKIKOIINMWQGVKAMYAPPINGOIRCSSNI	448
Oy	348	TGILLTTDDGGINENGETELFRFGGDMDNMWSSELTKRVVKIEPLEGVAFTCKKRYVORE	407
Dd	449	TGILLTTDDGGSRSREIFPRGGGMNMDNMSSELTKRVVKIEPLGVAFTAARRVORE	508
Oy	408	KRAVG--GAVALGFAGAAGSTMGAASMTLTVOQALLSGIVOOQNNLIRATEAOORMLQL	466
Dd	509	KRAVGAIGAMEFLGFLGAAGSTMGAASLTLYVOAQOLSGIVOOQNNLIRALEAOOHLOQL	568
Oy	467	TVMGIKQLQAVLAVERYLDQDLLGIMGCSGKLICTTAVPWNASWSNKSILDRITNNMTW	526
Dd	569	TVMGIKQLQAVLAVERYLRDQDLLGIMGCSGKLICTTTPVNANWSNKSIXSDQIKNNMTW	628
Oy	527	MEMEREIDNTYSLEYTLIESONOQEKNQDELBELDKASIAMNFDTTNMLMY	579
Dd	629	MEMEREIDNTYSLEYTLIESONOQEKNQDELBELDKASIAMNFSITNMLMY	681

RESULT 8
ENV_HV1PV STANDARD; PRT; 856 AA.

AC PO3376;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-1993 (Rel. 38, Last annotation update)
DE Envelope polypeptidein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].

GN ENV.
 OS Human immunodeficiency virus type 1 (PV22 isolate) (HIV-1).
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
 NCBI_TaxID=11700;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=8511157; PubMed=2982104;
 RA Mesling M.A., Smith D.H., Cabradilla C.D., Benton C.V., Laesky L.A.,
 "Nucleic acid structure and expression of the human
 RT AIDS/lymphadenopathy retrovirus.",
 RL Nature 313:450-458(1985).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: K02083; AAB59873.1; -.
 DR EMBL: X01762; CAA25903.1; ALT_SEQ.
 DR PIR: A03974; VCLJTV.
 DR HIV: K02083; ENVSPV22.
 DR InterPro: IPR000328; Env GP41.
 DR InterPro: IPR000777; GP120.
 DR Pfam: PF00516; GP120; 1.
 DR Pfam: PF00517; GP41; 1.
 KM AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
 KM Signal.
 FT SIGNAL. 1 30
 FT CHAIN 511 511 EXTERIOR MEMBRANE GLYCOPROTEIN.
 FT DISULFID 54 74 TRANSMEMBRANE GLYCOPROTEIN.
 FT DISULFID 119 205 BY SIMILARITY.
 FT DISULFID 126 196 BY SIMILARITY.
 FT DISULFID 131 157 BY SIMILARITY.
 FT DISULFID 218 247 BY SIMILARITY.
 FT DISULFID 228 239 BY SIMILARITY.
 FT DISULFID 296 331 BY SIMILARITY.
 FT DISULFID 378 445 BY SIMILARITY.
 FT DISULFID 385 418 BY SIMILARITY.
 FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 625 625 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 856 AA; 97339 MW; 5FCDB1DC3C1209B3 CRC64;
 Query Match 86.8%; Score 2715; DB 1; Length 856;
 Best Local Similarity 79.9%; Pred. No. 5e-212;
 Matches 520; Conservative 27; Mismatches 30; Indels 74; Gaps 6;
 QY 2 EKLWVTVYGVVWPKKATTTLPFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENVTE 61
 DB 32 EKLWVTVYGVVWPKKATTTLPFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENVTE 91
 QY 62 HFNWKKNNWTEQOEEDIIISLMDOSLKPCKVLTPLCGA----- 98
 DB 92 NFNWKKNNWTEQOEEDIIISLMDOSLKPCKVLTPLCGA----- 98
 QY 99 -----GCDTSVTOACPKISFE 115
 DB 152 GEIKKCSFNISTIRGKQVEYAFPYKLDIIPIDNDTSTYLTSCNTSVITQCPKVSFE 211
 QY 116 PIPHYCAPAGFAILKCNDRTPNGKGPCKNVSTVOCTHGIRPVSTOLLNGSLAEDEV 175
 DB 212 PIPHYCAPAGFAILKCNDRTPNGKGPCKNVSTVOCTHGIRPVSTOLLNGSLAEDEV 271
 QY 176 IRSDNFTNNAKTTIVOLKESVEINCTRPNNTRKSHI--GPGAPYTTSEITGDIRQAH 233
 DB 272 IRSANFTDNAKTTIVOLKESVEINCTRPNNTRKSHI--GPGAPYTTSEITGDIRQAH 330
 QY 234 CNISRAKMNNTLKOIVTKLEOF--ENKTIVFHNSSGGDEPIVHNSFGGSEFPYCNSTOL 292
 DB 331 CNISRAKMNNTLKOIVTKLEOF--ENKTIVFHNSSGGDEPIVHNSFGGSEFPYCNSTOL 390
 QY 293 FNSFWNN---NTEGSNNTG--NTITLPCRIKQIINMWQVGMAYAPPIRGQIRCSNIT 348
 DB 391 FNSFWNN---NTEGSNNTG--NTITLPCRIKQIINMWQVGMAYAPPIRGQIRCSNIT 450
 QY 349 GLITTRPGGNGENGTIFRPGCGDMRDMRSELYKRVKVIKELGVAFTCKRRVVOREK 408
 DB 451 GLITTRPGGNGENGTIFRPGCGDMRDMRSELYKRVKVIKELGVAFTCKRRVVOREK 510
 QY 409 RAVGIGAVPLGFGAAGSTGGAASMTLTVOARLLSGIVQOQNNLRAIEAQQMLQTLV 468
 DB 511 RAVGIGAVPLGFGAAGSTGGAASMTLTVOARLLSGIVQOQNNLRAIEAQQMLQTLV 570
 QY 469 WGIKQARVLAVERVYGDQOLLGIVGCSGKLTCTAVPNVNASMSKSLDRINWNTWME 528
 DB 571 WGIKQARVLAVERVYGDQOLLGIVGCSGKLTCTAVPNVNASMSKSLDRINWNTWME 630
 QY 529 WEREIDNYTSEIYTLIESQOQEKQOELELDKMASLWNFDTWMLWY 579
 DB 631 WEREIDNYTSEIYTLIESQOQEKQOELELDKMASLWNFDTWMLWY 681
 RESULT 9
 ENV_HV1A2 STANDARD; PRT; 855 AA.
 ID ENV_HV1A2
 AC P03378;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Envelope polyprotein GP160 precursor [containing: Exterior membrane
 glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
 GN ENV.
 OS Human immunodeficiency virus type 1 (ARV2/SF2 isolate) (HIV-1).
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
 NCBI_TaxID=11685;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85090453; PubMed=2578227;
 RA Sanchez-Pescador R., Power M.D., Barr P.J., Steimer K.S.,
 Stempien M.M., Brown-Shimer S.L., Gee W.W., Renard A., Randolph A.,
 Levy J.A., Dina D., Luciw P.A.,
 "Nucleotide sequence and expression of an AIDS-associated retrovirus
 RT (ARV-2).";

RL Science 227:484-492(1985).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: K02007; AAB59882.1; -
 DR PIR: A03976; VCLA2.
 DR HIV: K02007; ENV5SF2.
 DR InterPro: IPR000328; Env GP41.
 DR InterPro: IPR000777; GP120.
 DR Pfam: PF00516; GP120; 1.
 DR Pfam: PF00517; GP41; 1.
 DR AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
 KM signal.
 KM SIGNAL 1 29
 CHAIN 30 509
 CHAIN 510 855
 FT DISULFID 53 73
 FT DISULFID 118 208
 FT DISULFID 125 199
 FT DISULFID 130 155
 FT DISULFID 221 250
 FT DISULFID 231 242
 FT DISULFID 289 333
 FT DISULFID 360 442
 FT DISULFID 387 415
 FT CARBOHYD 87 87
 FT CARBOHYD 129 129
 FT CARBOHYD 140 140
 FT CARBOHYD 154 154
 FT CARBOHYD 158 158
 FT CARBOHYD 184 184
 FT CARBOHYD 190 190
 FT CARBOHYD 200 200
 FT CARBOHYD 223 233
 FT CARBOHYD 244 244
 FT CARBOHYD 265 265
 FT CARBOHYD 279 279
 FT CARBOHYD 292 292
 FT CARBOHYD 298 298
 FT CARBOHYD 304 304
 FT CARBOHYD 334 334
 FT CARBOHYD 341 341
 FT CARBOHYD 358 358
 FT CARBOHYD 364 364
 FT CARBOHYD 388 388
 FT CARBOHYD 394 394
 FT CARBOHYD 400 400
 FT CARBOHYD 408 408
 FT CARBOHYD 445 445
 FT CARBOHYD 458 458
 FT CARBOHYD 461 461
 FT CARBOHYD 610 610
 FT CARBOHYD 615 615
 FT CARBOHYD 624 624
 FT CARBOHYD 636 636
 FT CARBOHYD 815 815
 SQ SEQUENCE 855 AA; 97438 MW; A3BC20573AAC1A2 CR664;
 Query Match 86.7%; Score 2714; DB 1; Length 855;
 Best Local Similarity 79.3%; Pred. No. 6; 1e-212;
 Matches 518; Conservative 27; Mismatches 30; Indels 78; Gaps 7;
 QY 2 EKLMTVYVGVVWKEATTTTLCASDARAYDTEVNNWATTAQVCTDDNPQVUENYTE 61
 DB 31 EKLMTVYVGVVWKEATTTTLCASDARAYDTEVNNWATTAQVCTDDNPQVUENYTE 90
 QY 62 HFNMMKNMVEOMEDIIISLDQSLKPCVKLTPLCGA----- 98

DB 91 NFNMMKNMVEOMEDIIISLDQSLKPCVKLTPLCTVLTNCTDLGKATNTSSNMKEIIG 150
 QY 69 -----GCDPSVTIOAPRK 111
 DB 151 EIKKCSFNITTSIRDKIOKENALFRNLDPVIDNASTTNTYRLIHCKNSVITIOAPRK 210
 QY 112 ISFEPPIPHYCAPAGAFALCKNDKTPNGKSPKAVSTVQCTHGIRPVVSTOLLGSLAE 171
 DB 211 VSFEPIPHYCTPAGFALCKNKTFRNGKSGCTNVSTVQCTHGIRPIVSTOLLGSLAE 270
 QY 172 EEWVIRSDNFTNNAKTIIVOLKESVEINCTRPNNTRKSIHIGPRAFTTGEIIGDIRQ 231
 DB 271 EEWVIRSDNFTNNAKTIIVOLKESVEINCTRPNNTRKSIHIGPRAFTTGEIIGDIRK 330
 QY 232 AHCNISRAKNDTLKOIYIKRBOF-EKTIIVFNHSSGDPDEIVWHSNCGSEFFYCNST 290
 DB 331 AHCNISRAQMNNTTEIOYIKLRBOFGNNKTIIVFNHSSGDPDEIVWHSNCGSEFFYCNST 390
 QY 291 QLFNSTWN-NNTGSGNTEGN-TITLPCRIOINMOEYKAMVAPPITRGQIRCSNIT 348
 DB 391 QLFNSTWNLNTEG---TKGDTIILPCRIOINMOEYKAMVAPPITRGQIRCSNIT 447
 QY 349 GLLLTRODGIN-ENGTEIFRPGGDMRDNMSSELYKYVXIEPLGVAPTYCKRRVQRE 407
 DB 448 GLLLTRODGTNVTNDETVFRPGGDMRDNMSSELYKYVXIEPLGVAPTYCKRRVQRE 507
 QY 408 KRAVCI-GAVLFGPGAAGSTMGASMTLTQARLLSGIYVOQNNLLRAIEAOQMLQL 466
 DB 508 KRAVGIIVGAMPLGFGAAGSTMGAVSLTLTQARQLLSGIYVOQNNLLRAIEAOQMLQL 567
 QY 467 TWNGIKOLQAVLAVERYLPGQOLLGIGWCSGSKLICCTAVPMNASWSKSLDRIINNMTW 526
 DB 568 TWNGIKOLQAVLAVERYLPGQOLLGIGWCSGSKLICCTAVPMNASWSKSLDRIINNMTW 627
 QY 527 MEMEREIDNYSSEIYTLLEESQNOEKXOELEBLDKWASLMNMFDTNNMLMY 579
 DB 628 MOMEERIDNNTYITLLEESQNOEKXOELEBLDKWASLMNMFDTNNMLMY 680
 RESULT 10
 ENV_HVIY2
 ID ENV_HVIY2 STANDARD; PRT; 843 AA.
 AC P35961;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Envelope glycoprotein gp160 precursor [Contains: Exterior membrane
 GN glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
 OS Human immunodeficiency virus type 1 (YU-2 isolate) (HIV-1).
 OC Viruses; Retroid viruses; Retroviridae; Lentiviridae.
 OX NCBI_TaxID=36377;
 RN [1]
 RP SEQUENCE FROM N. A.
 RX MEDLINE=93021387; PubMed=1404605;
 RA Li Y., Hui H., Burgess C.J., Price R.W., Sharp P.M., Hahn B.H.,
 RT Shaw G.M.;
 RT "Complete nucleotide sequence, genome organization, and biological
 RT properties of human immunodeficiency virus type 1 in vivo: evidence
 RT for limited defectiveness and complementation.";
 RL J. Virol. 66:6587-6600(1992).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M93258; -; NOT_ANNOTATED_CDS.
 DR PIR: H44001; H44001.


```

DR PDB; 1G9N; 27-DEC-00.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00517; GP41; 1.
DR Pfam; PF00516; GP120; 1.
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal;
KW 3D-structure.
FT SIGNAL 1 29
FT CHAIN 30 489 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT TRANSMEM 490 843 TRANSMEMBRANE GLYCOPROTEIN.
FT TRANSMEM 738 755 POTENTIAL.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 201 BY SIMILARITY.
FT DISULFID 125 192 BY SIMILARITY.
FT DISULFID 130 155 BY SIMILARITY.
FT DISULFID 214 243 BY SIMILARITY.
FT DISULFID 224 235 BY SIMILARITY.
FT DISULFID 292 326 BY SIMILARITY.
FT DISULFID 373 432 BY SIMILARITY.
FT DISULFID 380 405 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 129 129 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 138 138 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 154 154 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 158 154 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 193 193 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 258 258 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 285 285 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 291 291 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 297 297 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 327 327 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 351 351 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 381 381 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 389 389 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 395 395 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 435 435 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 450 450 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 598 598 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 603 603 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 612 612 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 803 803 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE 843 AA; 95648 MM; C69DFD971C918B71 CRC64;

Query March 86.5%; Score 2706; DB 1; Length 843;
Best Local Similarity 79.9%; Pred. No. 2.7e-211;
Matches 513; Conservative 25; Mismatches 36; Indels 68; Gaps 5;

```

```

QY 239 AKNDTLKQIVIKLRQF-ENKTIIVFNHSSGDPPEIWMHSFNGCGEEFYCNSTOLFNSTW 297
DQ 331 TQMENTLEQALIKLKEKFGNKKTIIRPSSGSGPPEIVTHSFNCGEEFYCNSTOLF--TW 308
QY 298 NNNTSGSNNTSGNTTTLPCRKOIINNMQEVGKAMAPPIRGOIRCGSSNTTGLLTRDGG 357
DQ 389 -NDRKLNT-GRNITLPCRKOIINNMQEVGKAMAPPIRGOIRCGSSNTTGLLTRDGG 446
QY 358 INENGTETFPGGDMDDNRSELVYKVKIEPLGAPTRCKRRVQREKRAVGAVF 417
DQ 447 KDTNGETIFRPGGDMDDNRSELVYKVKIEPLGAPTRCKRRVQREKRAVGAVF 506
QY 418 LGFLGAGSTMGASMTLTVQARLLSGIVQOQNNLLRAIEAOQRLQTLTWGIKQLOAR 477
DQ 507 LGFLGAGSTMGASMTLTVQARLLSGIVQOQNNLLRAIEAOQRLQTLTWGIKQLOAR 566
QY 478 VLAVERYLDQQLIGWCGSKLICCTAVPMNASMSKSIDRIWNNNTMWEREIDNYT 537
DQ 567 VLAVERYLDQQLIGWCGSKLICCTAVPMNASMSKSIDRIWNNNTMWEREIDNYT 626
QY 538 SEIYTLLESONOQEKKEOELLKDKASLMMNPDLTNWIMY 579
DQ 627 HIYSLIEGSONOQEKKEOELLKDKASLMMNPDLTNWIMY 668

RESULT 11
ENV_HV1H2 STANDARD; PRT; 856 AA.
ID ENV_HV1H2
AC P04578; 009779;
DT 13-AUG-1987 (Rel. 05, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (HXB2 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_Taxid=11706;
OX [1]
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87299196; Pubmed=3040055;
RA Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,
RA Gallo R.C., Wong-Staal F.;
RT "Complete nucleotide sequences of functional clones of the AIDS
virus ";
RL AIDS Res. Hum. Retroviruses 3:57-69(1987).
RN [2]
RP REVISIONS.
RA Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,
RA Gallo R.C., Wong-Staal F.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K03455; AAB50262.1; -
DR EMBL; AF038399; AAB99976.1; -
DR EMBL; AF033819; AAC82596.1; -
DR PDB; 1DF4; 26-JAN-00.
DR PDB; 1DF5; 26-JAN-00.
DR PDB; 1DLB; 02-APR-00.
DR PDB; 1G9M; 27-DEC-00.
DR PDB; 1GCI; 19-AUG-98.
DR PDB; 1GZL; 10-OCT-02.
DR PDB; 1K3J; 10-OCT-01.
DR PDB; 1K34; 10-OCT-01.
DR HIV; K03455; ENVSHXB2.

```



```

DR InterPro: IPR000328; Env_Gp41.
DR InterPro: IPR000777; Gp120.
DR Pfam: PF00516; Gp120; 1.
DR Pfam: PF00517; Gp41; 1.
KW Aids; Coat protein; Glycoprotein; Transmembrane; Signal;
KW 3D-structure.
FT SIGNAL 1 30
FT CHAIN 31 511
FT CHAIN 512 856
FT DISULFID 54 74
FT DISULFID 119 205
FT DISULFID 126 196
FT DISULFID 131 157
FT DISULFID 218 247
FT DISULFID 228 239
FT DISULFID 296 331
FT DISULFID 378 445
FT DISULFID 385 418
FT CARBOHYD 88
FT CARBOHYD 136
FT CARBOHYD 141
FT CARBOHYD 156
FT CARBOHYD 160
FT CARBOHYD 166
FT CARBOHYD 187
FT CARBOHYD 197
FT CARBOHYD 230
FT CARBOHYD 234
FT CARBOHYD 241
FT CARBOHYD 262
FT CARBOHYD 276
FT CARBOHYD 289
FT CARBOHYD 295
FT CARBOHYD 301
FT CARBOHYD 312
FT CARBOHYD 332
FT CARBOHYD 339
FT CARBOHYD 356
FT CARBOHYD 366
FT CARBOHYD 392
FT CARBOHYD 397
FT CARBOHYD 406
FT CARBOHYD 448
FT CARBOHYD 463
FT CARBOHYD 611
FT CARBOHYD 616
FT CARBOHYD 624
FT CARBOHYD 637
FT CARBOHYD 674
FT CARBOHYD 750
FT CARBOHYD 816
SEQUENCE 856 AA; 97212 MW; 6FAB16AF85107FE0 CRC64;

Query Match
Beet Local Similarity 79.6%; Score 2702; DB 1; Length 856;
Matches 518; Conservative 28; Mismatches 31; Indels 74; Gaps 6;

2 EKLWTVVYGVVPMKEATTTLCASDAPAKYDTEVHVWATHACVPTDPNPOEVLNVNVE 61
32 EKLWTVVYGVVPMKEATTTLCASDAPAKYDTEVHVWATHACVPTDPNPOEVLNVNVE 91
62 HENMKANNVQEMODIISLMDQSLKPCYKLTPLCGA----- 98
92 NFNMMKNDVQEMHEDIISLMDQSLKPCYKLTPLCGSLKCTDLKNDTNTSSGRIMK 151
99 -----GCDTSYITQACPKLSE 115
152 GEIKNCSFNISTIRGKQKEAYAFYKLDIIPIDNNTSYKLTSCNTSYITQACPKVSP 211
116 PIPHYCAPAGFAILKNDKTFNGKPCGNVSTVOCTHGRIPVSTOLLNGLAEEVY 175
212 PIPHYCAPAGFAILKNDKTFNGKPCGNVSTVOCTHGRIPVSTOLLNGLAEEVY 271
176 IISDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIH--GPGRAFTTGEIIGDIRAH 233

```

```

Db 272 IRSVNFNDNAKTIIVQLNVTSEINCTRPNNNTRKRIORGRAFTVIGK-IGNMROAH 330
Qy 234 CNISRAKNMDTLKQIVIKLREOF-ENKTIIVFHSSGDEPIYHSPNCGEFPYCNSTOL 292
Db 331 CNISRAKNMDTLKQIVIKLREOF-ENKTIIVFHSSGDEPIYHSPNCGEFPYCNSTOL 390
Qy 293 FNSSTWNN--NTGSSNTEG-NTITLPCRIRKQIIMMDEVGKAMYAPPIRGQIRCSNIT 348
Db 391 FNSSTWNN--NTGSSNTEG-NTITLPCRIRKQIIMMDEVGKAMYAPPIRGQIRCSNIT 450
Qy 349 GLLLRDQGINENGRIEIRPGGDMRDNRSLKYKVKIPLGVAAPFKCARVYOREK 408
Db 451 GLLLRDQGINENGRIEIRPGGDMRDNRSLKYKVKIPLGVAAPFKCARVYOREK 510
Qy 409 RAVGIGAVEFLGAGASTMGAASTMTLVQARLLSGIYQOQNNLIRALEAOQHLQITV 468
Db 511 RAVGIGAVEFLGAGASTMGAASTMTLVQARLLSGIYQOQNNLIRALEAOQHLQITV 570
Qy 469 WGIKQLQARVLAVERYLGDQDLIGGSGKLICTTAVPMNMSWNSKSLDRIMNNMTWNE 528
Db 571 WGIKQLQARVLAVERYLGDQDLIGGSGKLICTTAVPMNMSWNSKSLDRIMNNMTWNE 630
Qy 529 WEREIDNTTSYITLIESONQOEYRDELLFELDKWASIMKFDITNMLWY 579
Db 631 WEREIDNTTSYITLIESONQOEYRDELLFELDKWASIMKFDITNMLWY 681

RESULT 12
ENV_HV1H3 STANDARD; PRT; 856 AA.
AC P04624;
DT 13-AUG-1987 (rel. 05, Created)
DT 01-FEB-1996 (rel. 33, Last sequence update)
DT 15-SEP-2003 (rel. 42, Last annotation update)
DE Envelope glycoprotein Gp160 precursor [Contains: Exterior membrane
glycoprotein (Gp120); Transmembrane glycoprotein (Gp41)].
GN ENV.
OS Human immunodeficiency virus type 1 (HXB3 isolate) (HIV-1).
OC Viruses; Retroviruses; Retroviridae; Lentiviruses.
OX NCBI_Taxid=11707;
RN [1]
RP MEDLINE=85228248; PubMed=2988795;
RA Crowl R., Ganguly K., Gordon M., Conroy R., Schaber M., Kramer R.,
Shaw G.M., Wong-Staal F., Reddy E.P.;
RT "HIV-1 env gene products synthesized in E. coli are recognized by
antibodies present in the sera of AIDS patients.";
RL Cell 41:979-986 (1985).
CC -----
CC This SMS-S-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib.ch).
CC -----
DR EMBL: M14100; AAA44679.1; -
DR PDB: 1DAU; 17-OCT-01.
DR PDB: 1DAV; 17-OCT-01.
DR HIV: M14100; ENVSHXB3.
DR InterPro: IPR000328; Env_Gp41.
DR InterPro: IPR000777; Gp120.
DR Pfam: PF00516; Gp120; 1.
DR Pfam: PF00517; Gp41; 1.
KW Aids; Coat protein; Glycoprotein; Transmembrane; Signal;
KW 3D-structure.
FT SIGNAL 1 30
FT CHAIN 31 511
FT CHAIN 512 856
FT DISULFID 119 205
FT DISULFID 126 196
FT DISULFID 131 157
FT DISULFID 218 247
FT DISULFID 228 239
FT DISULFID 296 331
FT DISULFID 378 445
FT DISULFID 385 418
FT CARBOHYD 88
FT CARBOHYD 136
FT CARBOHYD 141
FT CARBOHYD 156
FT CARBOHYD 160
FT CARBOHYD 166
FT CARBOHYD 187
FT CARBOHYD 197
FT CARBOHYD 230
FT CARBOHYD 234
FT CARBOHYD 241
FT CARBOHYD 262
FT CARBOHYD 276
FT CARBOHYD 289
FT CARBOHYD 295
FT CARBOHYD 301
FT CARBOHYD 312
FT CARBOHYD 332
FT CARBOHYD 339
FT CARBOHYD 356
FT CARBOHYD 366
FT CARBOHYD 392
FT CARBOHYD 397
FT CARBOHYD 406
FT CARBOHYD 448
FT CARBOHYD 463
FT CARBOHYD 611
FT CARBOHYD 616
FT CARBOHYD 624
FT CARBOHYD 637
FT CARBOHYD 674
FT CARBOHYD 750
FT CARBOHYD 816
SEQUENCE 856 AA; 97212 MW; 6FAB16AF85107FE0 CRC64;

```


FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 856 AA; 96938 MW; 0C241332CF7E6687 CRC64;

Query Match

86.1%; Score 2693; DB 1; Length 856;

Best Local Similarity 79.1%; Pred. No. 3.1e-210;

Matches 515; Conservative 29; Mismatches 33; Indels 74; Gaps 6;

QY 2 EKLMTVTVYGVVPMKEATITTLFCASDAKAYDTEVNNVWATHACVPTDNPQEVLENTE 61
 DB 32 EKLMTVTVYGVVPMKEATITTLFCASDAKAYDTEVNNVWATHACVPTDNPQEVLENTE 91
 QY 62 HNNMMKNNNVVEQMOEDIIISLMDPSLKPCYKLTPLGCA----- 98
 DB 92 HNNMMKNNDVVEQMHEDIISLMDPSLKPCYKLTPLCVSLKCTDKNDNTNNSSGGIMEX 151
 QY 99 -----GCDTSVITQACPKEISPE 115
 DB 152 GEIKNCSFNISTIRGKVOKEVAFYKHDIIPIDNDTSTYTLSCNTSVITQACPKEISPE 211
 QY 116 PIPHYCAPAGFAILLKNDKTFNGKGPCQNVSTVOCTHGIRPVSTQLNLNGSLAEVY 175
 DB 212 PIPHYCAPAGFAILLKNDKTFNGKGPCQNVSTVOCTHGIRPVSTQLNLNGSLAEVY 271
 QY 176 IRSDNFTNAKTIIVQLKESVEINCTRPNNNTRKSIH--GPRAFYTTGELIIGDIRAH 233
 DB 272 IRSANLTDVKTIIIVQLNOSVEINCTRPNNNTRKRIORGPRTFTYIGK--IGNRORAH 330
 QY 234 CNISRAKNNDTLKOIIVIKLREOF-ENKTIIVFNHSSGDEPEIVHSHFNCGEFFYCNSYOL 292
 DB 331 CNISRAKNNDTLKOIIVIKLREOF-ENKTIIVFNHSSGDEPEIVHSHFNCGEFFYCNSYOL 390
 QY 293 FNSTNN--NTEGSSNTEG-NTTILPCRIKOIINMMOEVGKAMTAPPIRGOIRCSSNIT 348
 DB 391 FNSTWFSNSTEGSSNTEGSDTITLPCRIKOIINMMOEVGKAMTAPPIRGOIRCSSNIT 450
 QY 349 GLLLTRDGINENGTEIFRPGGDMRDNRSELYKKVKKIEPLGVAPTKCRVAVOREK 408
 DB 451 GLLLTRDGGNNNGSEIFRPGGDMRDNRSELYKKVKKIEPLGVAPTKCRVAVOREK 510
 QY 409 RAVGIGAVELGFLGAAGSTMGASMTLTVQARLLLSGIVOQONNLRAIEAOQRLQLTV 468
 DB 511 RAVGIGAVELGFLGAAGSTMGASMTLTVQARLLLSGIVOQONNLRAIEAOQRLQLTV 570
 QY 468 WGIKQLOARVLAVERYLGDOQLGIWCGGKLICTTAVPWNASWSNKSIDRIWNNNTWME 528
 DB 571 WGIKQLOARVLAVERYLGDOQLGIWCGGKLICTTAVPWNASWSNKSIDRIWNNNTWME 630
 QY 529 WREIDNTYSEIYTLIEESONOEKNEBELLDKASLMMNPDTNMLMY 579
 DB 631 WREIDNTYSEIYTLIEESONOEKNEBELLDKASLMMNPDTNMLMY 681

Search completed: December 12, 2003, 12:30:48
 Job time : 11.7179 secs

THIS PAGE BLANK (US:70)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: December 12, 2003, 12:24:08 ; Search time 37.2225 Seconds

(without alignments)
2665.163 Million cell updates/sec

Title: US-10-032-162-17

Perfect score: 3351
Sequence: 1 VEKLMVTYYGVVPMKEATL.....EIDKWSASIMNFDITKMLMY 625

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

1 number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A.GeneSeq_19Jun03:*

1:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT:*
2:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT:*
3:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT:*
4:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT:*
5:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT:*
6:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT:*
7:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT:*
8:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT:*
9:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT:*
10:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT:*
11:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT:*
12:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT:*
13:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT:*
14:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT:*
15:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT:*
16:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT:*
17:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT:*
18:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT:*
19:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT:*
20:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT:*
21:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT:*
22:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT:*
23:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT:*
24:	/SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3345	99.8	625	AAAB61507	HIV-1 delTAV3 SOS
2	3295	98.3	643	AAAB61505	HIV-1 SOS gp140 g1
3	2968.5	88.6	850	AAAB67724	gp120 from the HIV
4	2965	88.5	883	AAAB83761	Ancestral HIV-1 gr
5	2938.5	87.7	855	AAAB1581	Human Immunodefici
6	2938.5	87.7	855	AAAB8113	Env protein of the
7	2928	87.4	847	AAI97073	Variant HIV-1 SF16
8	2923	87.2	579	AAAB61506	HIV-1 delTAV12* S
9	2923	87.2	24	ABU65655	Human immunodefici

10	2918	87.1	842	23	ABBB06211
11	2898.5	86.5	868	27	AAAB60422
12	2895.5	86.4	868	23	AAO13389
13	2894.5	86.4	868	7	AAAB60063
14	2887	86.2	856	22	AAAB85999
15	2887	86.2	863	14	AAAB43869
16	2886	86.1	854	21	AAAB10697
17	2886	86.1	854	21	AAAB10053
18	2886	86.1	854	22	AAAB86199
19	2886	86.1	854	22	AAAB67277
20	2886	86.1	854	23	ABG73663
21	2885	86.1	856	7	AAAB60131
22	2884.5	86.1	851	7	AAAB80967
23	2883	86.0	856	7	AAAB61514
24	2883	86.0	856	20	AAAB89325
25	2881.5	86.0	901	8	AAAB70665
26	2879	85.9	856	13	AAAB25940
27	2878	85.9	856	21	AAAB97072
28	2877	85.9	863	7	AAAB60349
29	2876	85.8	857	16	AAAB67725
30	2871	85.7	863	14	AAAB43873
31	2871	85.7	1188	17	AAAB94622
32	2870	85.6	856	18	AAAB1579
33	2870	85.6	856	20	AAAB8111
34	2869.5	85.6	751	10	AAAB90604
35	2867	85.6	856	22	AAAB45697
36	2861	85.4	857	12	AAAB49904
37	2861	85.4	857	18	AAAB1580
38	2861	85.4	857	20	AAAB8112
39	2858.5	85.3	855	19	AAAB43069
40	2854.5	85.2	880	18	AAAB23333
41	2854.5	85.2	880	20	AAAB73332
42	2854.5	85.2	880	21	AAAB14844
43	2854	85.2	858	19	AAAB43067
44	2853	85.1	856	20	AAAB90178
45	2851.5	85.1	855	12	AAAB14905

ALIGNMENTS

RESULT 1					
AAAB61507					
ID	AAAB61507	standard; Protein; 625 AA.			
AC	AAAB61507;				
DT	05-APR-2001	(first entry)			
DE	HIV-1 delTAV3 SOS gp140 glycoprotein.				
KM	gp120; anti-HIV; gene therapy; vaccine; viral envelope protein; gp141;				
KW	gp140.				
OS	Human immunodeficiency virus type 1.				
PN	WO200100648-A1.				
PD	04-JAN-2001.				
PF	23-JUN-2000; 2000WO-US17267.				
PR	25-JUN-1999; 99US-0340992.				
PA	(PROG-) PROGENICS PHARM INC.				
PA	(AARO-) AARON DIMOND AIDS RES CENT.				
PI	Binley JM, Schuelke N, Olson WC, Madden PJ, Moore JP;				
DR	WPI; 2001-122993/13.				
DR	N-PSDB; AAF28583.				
PT	New viral envelope proteins, useful for producing vaccines to treat				

HIV Env isolate SF
Sequence of LAV vi
Lymphadenopathy-as
HIV virus env gene
Amino acid sequenc
HIV-III ENV-LOR g
HIV-1 env protein.
HIV-1 coat protein
HIV gp41 DNA SEQ I
Protein encoded by
HIV-1 NL4-3 gp120
Sequence of the A1
HIV protein HT6.
Sequence of envlo
HIV-1 env protein
Sequence encoded b
Modified HIV env g
Wild type HIV-1 HX
HIV-III virus (HI
gp120 from the HIV
HIV-III ENV-LOR g
HIV-1 gp160-NY5 p5
Human Immunodefici
Predicted amino ac
HIV-1 env segment
HIV-1/IIIB env clo
HIV-1 (MN-ST1) env
Human Immunodefici
Predicted amino ac
HIV-1 gp120 protei
Human immunodefici
HIV envelope prote
HIV envelope prote
HIV-1 gp120 protei
HIV-III env prote
HIV-1 BA-L clone-e

PT human immunodeficiency virus-1 infections, comprises amino acid
 PT sequence mutations such that viral transmembrane-surface protein
 PT complex is more stable
 PS Disclosure; Fig 15; 109pp; English.
 CC The present invention relates to a viral envelope protein. The viral
 CC envelope protein comprises a viral surface protein (e.g. glycoprotein
 CC gp120) and a corresponding viral transmembrane protein (e.g. gp41), in
 CC which the viral envelope protein contains one or more amino acid sequence
 CC mutations that enhance the stability of the complex formed between the
 CC viral surface and transmembrane proteins. The viral envelope protein can
 CC be used in the treatment of viral infection e.g. HIV-1 infection. The
 CC present sequence is HIV-1 deltaV3 SOS gp140, which was used in the
 CC present invention.
 XX Sequence 625 AA;
 SQ
 - Query Match 99.8%; Score 3345; DB 22; Length 625;
 - Local Similarity 100.0%; Pred. No. 2.5e-188;
 Matches 625; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEKLMVTYYYGVPMWEKATTTLCASDAKAYDTEVHNWMAHACVPTDPNPOEVLENT 60
 DB 1 VEKLMVTYYYGVPMWEKATTTLCASDAKAYDTEVHNWMAHACVPTDPNPOEVLENT 60
 QY 61 EHFMMKNNWVEQOEEDIIISLMDQSLKPCVYKLTPLCYTLNCKDVNATNTNDSGTMERG 120
 DB 61 EHFMMKNNWVEQOEEDIIISLMDQSLKPCVYKLTPLCYTLNCKDVNATNTNDSGTMERG 120
 QY 121 EIKKCSFNITTSIRDEVOKEVALFYKLDVYXIDNNNTSYRLISCDTSVITQACPKISFEP 180
 DB 121 EIKKCSFNITTSIRDEVOKEVALFYKLDVYXIDNNNTSYRLISCDTSVITQACPKISFEP 180
 QY 181 IPIHYCAPAGFALLKCDKDTFNGKPKCNVSTXQCTHGIRPVVSTOLLNGSLAEEVVI 240
 DB 181 IPIHYCAPAGFALLKCDKDTFNGKPKCNVSTXQCTHGIRPVVSTOLLNGSLAEEVVI 240
 QY 241 RSDNFTNNAKTIIVOLKESVEINCTRPNNNAGDIRQAHNISRAKMDITLKOIVIKLRE 300
 DB 241 RSDNFTNNAKTIIVOLKESVEINCTRPNNNAGDIRQAHNISRAKMDITLKOIVIKLRE 300
 QY 301 QFENKTIIVFNHSSGGDPEIIVHSSFNCGEFPYCNSSTOLFNSNTNNTGSGNTITTL 360
 DB 301 QFENKTIIVFNHSSGGDPEIIVHSSFNCGEFPYCNSSTOLFNSNTNNTGSGNTITTL 360
 QY 361 PCRIKOIINMQEVEGKAMYPAPPIRGQIRCSNITGLLTRDGINENGTETIFRPGGDMR 420
 DB 361 PCRIKOIINMQEVEGKAMYPAPPIRGQIRCSNITGLLTRDGINENGTETIFRPGGDMR 420
 QY 421 DNMRSELKYRVVPIEPLGVAPTCKRRVVOREKRAVGIGAVPLGFGAAGSTMGAASMT 480
 DB 421 DNMRSELKYRVVPIEPLGVAPTCKRRVVOREKRAVGIGAVPLGFGAAGSTMGAASMT 480
 QY 481 LTVQARLLSGIIVQOONNLRAIEAQOQMLQLTWGIKQIOLARVLAVERVYGDQOLLGIW 540
 DB 481 LTVQARLLSGIIVQOONNLRAIEAQOQMLQLTWGIKQIOLARVLAVERVYGDQOLLGIW 540
 QY 541 GCSGKLICTTAVPNVNASWNSKSLDRINNMNTWMEWEREIDVYSEIYTLIEESQNOQCKN 600
 DB 541 GCSGKLICTTAVPNVNASWNSKSLDRINNMNTWMEWEREIDVYSEIYTLIEESQNOQCKN 600
 QY 601 EOELELDKASLNMWPDITKMWY 625
 DB 601 EOELELDKASLNMWPDITKMWY 625
 QY 601 EOELELDKASLNMWPDITKMWY 625
 DB 601 EOELELDKASLNMWPDITKMWY 625

RESULT 2
 AAB61505
 ID AAB61505 standard; Protein; 643 AA.
 XX AAB61505;
 XX

DT 05-APR-2001 (first entry)
 XX
 DE HIV-1 SOS gp140 glycoprotein.
 XX
 KM gp120; anti-HIV; gene therapy; vaccine; viral envelope protein; gp41;
 KM gp140.
 XX
 OS Human immunodeficiency virus type 1.
 XX
 PN W02001.00648-A1.
 XX
 PD 04-JAN-2001.
 XX
 PF 23-JUN-2000; 2000MO-US17267.
 XX
 PR 25-JUN-1999; 99US-0340992.
 XX
 PA (PROG-) PROGENICS PHARM INC.
 PA (AARO-) AARON DIAMOND AIDS RES CENT.
 XX
 PI Binley JM, Schelke N, Olson WC, Madden PJ, Moore JP;
 XX
 DR WPI; 2001-122993/13.
 DR N-PSDB; AAF28581.
 XX
 PT New viral envelope proteins, useful for producing vaccines to treat
 PT human immunodeficiency virus-1 infections, comprises amino acid
 PT sequence mutations such that viral transmembrane-surface protein
 PT complex is more stable
 PS Disclosure; Fig 13; 109pp; English.
 XX
 CC The present invention relates to a viral envelope protein. The viral
 CC envelope protein comprises a viral surface protein (e.g. glycoprotein
 CC gp120) and a corresponding viral transmembrane protein (e.g. gp41), in
 CC which the viral envelope protein contains one or more amino acid sequence
 CC mutations that enhance the stability of the complex formed between the
 CC viral surface and transmembrane proteins. The viral envelope protein can
 CC be used in the treatment of viral infection e.g. HIV-1 infection. The
 CC present sequence is HIV-1 SOS gp140, which was used in the present
 CC invention.
 XX
 SQ Sequence 643 AA;
 - Query Match 98.3%; Score 3295; DB 22; Length 643;
 - Local Similarity 96.0%; Pred. No. 2.2e-185;
 Matches 617; Conservative 0; Mismatches 8; Indels 18; Gaps 1;

QY 1 VEKLMVTYYYGVPMWEKATTTLCASDAKAYDTEVHNWMAHACVPTDPNPOEVLENT 60
 DB 1 VEKLMVTYYYGVPMWEKATTTLCASDAKAYDTEVHNWMAHACVPTDPNPOEVLENT 60
 QY 61 EHFMMKNNWVEQOEEDIIISLMDQSLKPCVYKLTPLCYTLNCKDVNATNTNDSGTMERG 120
 DB 61 EHFMMKNNWVEQOEEDIIISLMDQSLKPCVYKLTPLCYTLNCKDVNATNTNDSGTMERG 120
 QY 121 EIKKCSFNITTSIRDEVOKEVALFYKLDVYXIDNNNTSYRLISCDTSVITQACPKISFEP 180
 DB 121 EIKKCSFNITTSIRDEVOKEVALFYKLDVYXIDNNNTSYRLISCDTSVITQACPKISFEP 180
 QY 181 IPIHYCAPAGFALLKCDKDTFNGKPKCNVSTXQCTHGIRPVVSTOLLNGSLAEEVVI 240
 DB 181 IPIHYCAPAGFALLKCDKDTFNGKPKCNVSTXQCTHGIRPVVSTOLLNGSLAEEVVI 240
 QY 241 RSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRQAHN 282
 DB 241 RSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRQAHN 282
 QY 283 SRKAMDITLKOIVIKLEQFENKTIIVFNHSSGGDPEIIVHSSFNCGEFPYCNSSTOLFNS 342
 DB 301 SRKAMDITLKOIVIKLEQFENKTIIVFNHSSGGDPEIIVHSSFNCGEFPYCNSSTOLFNS 360
 QY 343 WNNNTGSGNTGNTITLPCRIKOIINMQEVEGKAMYPAPPIRGQIRCSNITGLLTRDG 402

```

Db      |||
361  WNNNTEGSSNTEGNTITTLPCR|KOIIMMOEVBKAMYPPIRGOIRCSNITGLLTRG 420
Qy      403  GINENGTEIFRPGGDMRDNRSELYKYVVKIEPLGVAPTCKRRVOREKRAVGIGAV 462
Db      421  GINENGTEIFRPGGDMRDNRSELYKYVVKIEPLGVAPTCKRRVOREKRAVGIGAV 480
Qy      463  FLGFLGAAGSTWGAASMTLTVQARLLLSGI VOOQNNLLRAIEAQRMLQLTWGIKQLQA 522
Db      481  FLGFLGAAGSTWGAASMTLTVQARLLLSGI VOOQNNLLRAIEAQRMLQLTWGIKQLQA 540
Qy      523  RVLAVERYIGDQQLGIGWCGSLICTTAVPNNASWSNKS LDRINNMNTMEMERIDNY 582
Db      541  RVLAVERYIGDQQLGIGWCGSLICTTAVPNNASWSNKS LDRINNMNTMEMERIDNY 600
Qy      583  TSEIYTLIEESONQOEKNEQELLEDKMASLNMWPDITKWLNY 625
Db      601  TSEIYTLIEESONQOEKNEQELLEDKMASLNMWPDITKWLNY 643

```

LT 3
5724

ID AAR67724 standard; Protein; 850 AA.

AC AAR67724;

DT 25-MAR-2003 (updated)

DT 07-SEP-1995 (first entry)

DE gp120 from the HIV GNE8 isolate.

KM HIV; human immunodeficiency virus; gp120; glycoprotein;

KW GNE clone; GNE8 isolate.

OS Human immunodeficiency virus type 1.

PN W09428929-A1.

PD 22-DEC-1994.

PF 07-JUN-1994; 94WO-US06036.

PR 07-JUN-1993; 93US-0072833.

PA (GETH) GENENTECH INC.

PI Berzhan PW, Nakamura GR;

XX WPI; 1995-036112/05.

XX N-PSDB; AAQ76018.

PT Use of HIV gp 120 polypeptide(s) - for developing probes for the

PT analysis, prevention and therapy of HIV infection

PS Claim 23; Page 27-30; 108pp; English.

CC This protein is the gp120 (envelope glycoprotein) from the GNE8 isolate
 CC of HIV-1 (human immunodeficiency virus type 1). There are neutralising
 CC epitopes in the V2 and C4 domains of gp120, in addition to the
 CC neutralising epitopes in the V3 domain. Although the amino acid sequences
 CC of the neutralising epitopes in the V2, V4 and C4 domains are variable,
 CC the amount of variation is highly constrained. This facilitates the
 CC design of HIV subunit vaccines that can induce antibodies that neutralise
 CC the most common HIV strains for a given geographic region. This invention
 CC provides a multivalent gp120 subunit vaccine where the gp120 present in
 CC the vaccine is from at least two HIV isolates which have different amino
 CC acid sequences for a neutralising epitope in these regions.
 CC (Updated on 25-MAR-2003 to correct PN field.)

SO Sequence 850 AA;

Query Match 88.6%; Score 2968.5; DB 16; Length 850;
 Best Local Similarity 86.9%; Pred. No. 3.8e-166;

Matches 563; Conservative 25; Mismatches 33; Indels 27; Gaps 6;

```

Qy      2  EKLAVTVYGVVPMKEATTLTFCASDAXAYDEVHNVNATACVPTDPPOVLENTYE 61
Db      31  EKLAVTVYGVVPMKEATTLTFCASDAXAYDEVHNVNATACVPTDPPOVLENTYE 90
Qy      62  HENMMKNMNVEMOBDIISLDOSLPCVKLTPLCVTLNCKDV-NATNTNDSGTMERG 120
Db      91  NFNMMKNMNVEMOBDIISLDOSLPCVKLTPLCVTLNCTDLKNATNTSSWGMERG 150
Qy      121  EIKNCSFNITTSIRBEVQKVALFYKLDVXVYIDNNNTSYRLISCTSVYTOACPKISFEP 180
Db      151  EIKNCSFNITTSIRBEVQKVALFYKLDVVPIDNNNTSYRLISCTSVYTOACPKISFEP 210
Qy      181  IPIHCAPAGFAILKCNKTFNGKPCKNVSTXOCTHGIRPVVSTOLLNGLABEEVYI 240
Db      211  IPIHCAPAGFAILKCRDKKNGTGPCTNVSTVQCTHGIRPVVSTOLLNGLABEEVYI 270
Qy      241  RSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRQAHANI 282
Db      271  RSANFSDNAKTIIVOLNESVEINCTRPNNNTRRSIHIGRAFYATGEIIGDIRQAHANI 330
Qy      283  SPAKNDPLKQIVITLREQFENKTIIVFNHSSGDEIYVHNSFCGGEFFYCNSTOLFNST 342
Db      331  SSTKMNNTLKQIVITLREHF-NKTIIVFNHSSGDEIYVHNSFCGGEFFYCNSTOLFNST 389
Qy      343  WN-----NTEGSSNTEGNTITTLPCR|KOIIMMOEVBKAMYPPIRGOIRCSNITGL 397
Db      390  WNTYTTWNTNTEGSSNT-GRNITLQCR|KOIIMMOEVBKAMYPPIRGOIRCSNITGL 448
Qy      398  LTRDGINENGTEIFRPGGDMRDNRSELYKYVVKIEPLGVAPTCKRRVOREKRAV 457
Db      449  LTRDGG-NNSETEIFRPGGDMRDNRSELYKYVVKIEPLGVAPTCKRRVOREKRAV 507
Qy      458  GIGAVFLGFLGAAGSTWGAASMTLTVQARLLLSGI VOOQNNLLRAIEAQRMLQLTWGI 517
Db      508  GIGAVFLGFLGAAGSTWGAASMTLTVQARLLLSGI VOOQNNLLRAIEAQRMLQLTWGI 567
Qy      518  KQIARVLAVERYLGDQQLGIGWCGSLICTTAVPNNASWSNKS LDRINNMNTMEMER 577
Db      568  KQIARVLAVERYLKDQQLGIGWCGSLICTTAVPNNASWSNKS LDKIWDNTMEMER 627
Qy      578  EIDNTSEIYTLIEESONQOEKNEQELLEDKMASLNMWPDITKWLNY 625
Db      628  EIDNTSLIYSLIEESONQOEKNEQELLEDKMASLNMWPDITKWLNY 675

```

RESULT 4

AAB82761 ID AAB82761 standard; Protein; 883 AA.

XX AAB82761;

DT 29-OCT-2001 (first entry)

DE Ancestral HIV-1 group M, subtype B gp160 protein.

KM HIV-1; env gene; gp160 protein; ancestral gene; phylogeny; vaccine.

OS Human immunodeficiency virus type 1.

PN W0200160838-AZ.

PD 23-AUG-2001.

PF 16-FEB-2001; 2001WO-US05288.

PR 18-FEB-2000; 2000US-0183659.

PA (UNITW) UNIV WASHINGTON.

XX Mullins JT, Rodrigo AG, Learn GH, Li F;

XX

Db 571 IKOLQARVLAVERYLKRDQQLIGWCGSGKLICTTAVPNNASMSKSLNKRIWNTWIEWD 630

Qy 577 REINDNTSEIYTLIEESONQOEKNEQELLELDKNASLWNMFDTIKMLWY 625

Db 631 REINNTSYIYSLIEESONQOEKNEQELLELDKNASLWNMFDTIKMLWY 679

RESULT 7

AA97073

AA97073 standard; protein; 847 AA.

AC AA97073;

DT 31-OCT-2000 (first entry)

XX

DE Variant HIV-1 SF162 Env gp160.

XX

KM HIV-1; vaccine; anti-HIV; immunogenic; modified; envelope; Env;

KM CD4 binding region; VI/V2 loop; bridging sheet.

XX

Human immunodeficiency virus type 1 isolate SF162.

XX

PN W0200039303-A2.

XX

PD 06-JUL-2000.

XX

PF 30-DEC-1999; 99WO-US31272.

XX

PR 31-DEC-1998; 98US-0114495.

PR 29-SEP-1999; 99US-0156670.

XX

PA (CHIR) CHIRON CORP.

PI Barnett S, Hartog K, Martin E;

XX

XX WPI; 2000-465745/40.

DR

XX

PT Novel modified HIV Env polypeptides useful as immunizing agents and for

PT preparing a vaccine to elicit an immune response against a broad range

PT of HIV subtypes

XX

XX Claim 5; Page 115-117; 139pp; English.

XX

CC Novel immunogenic modified human immunodeficiency virus (HIV) envelope

CC (Env) polypeptides having an amino acid deleted or replaced in the region

CC corresponding to residues 420-436 or 119-123 and 199-210 relative to the

CC isolate HXB-2 are disclosed. The modified Env polypeptide is based on HIV

CC strain SF162, with numbering relative to isolate HXB-2. The Env

CC polypeptides are modified so as to expose at least part of the CD4

CC binding region. The modified HIV Env polypeptides, coding polynucleotides

CC and constructs, further comprising an adjuvant, are used for inducing an

CC immune response in an individual. The method involves administering a

CC first composition comprising a polynucleotide encoding the Env

CC polypeptide in a priming step and administering a second composition

CC comprising a modified Env polypeptide as a booster in an amount

CC sufficient to induce an immune response in the individual. The first

CC and/or second composition further comprises an adjuvant (claimed). The

CC intracellularly produced Env polypeptides can be used for a number of

CC diagnostic and therapeutic purposes to determine the presence of reactive

CC antibodies/and or Env proteins in a biological sample to aid in the

CC diagnosis of HIV infection or disease status or as measure of response to

CC immunization.

XX

SO Sequence 847 AA;

Qy Query Match 87.4%; Score 2928; DB 21; Length 847;

Db Best Local Similarity 86.0%; Pred. No. 9e-164;

Matches 555; Conservative 28; Mismatches 40; Indels 22; Gaps 5;

Qy 1 VEKLMVTYYGVVPWKKEATITLFCASDAKAYDTEVHNWATACVPTPNQEVLEWNT 60

Db 30 VEKLMVTYYGVVPWKKEATITLFCASDAKAYDTEVHNWATACVPTPNQEVLEWNT 89

Qy 61 EHEFMWNNWQVQOEIIISLMDQSLKPCVKLTPLCVTLNCKOV-NATNTNDSGTMER 119

Db 90 ENFMWNNWQVQOEIIISLMDQSLKPCVKLTPLCVTLNCKOV-NATNTNDSGTMER 149

Qy 120 GEINCSFNITTSIRDEVQKEXALFYLDVYXIDNNNTSYRLISCDISVITQACPKISFE 179

Db 150 GEINCSFNITTSIRDEVQKEXALFYLDVYXIDNNNTSYRLISCDISVITQACPKISFE 209

Qy 180 PIPHYCAPAGFALIKCNDTFNGKXPCXNVSTQCTHGIRPVYSTOLLNGSLAESEV 239

Db 210 PIPHYCAPAGFALIKCNDTFNGKXPCXNVSTQCTHGIRPVYSTOLLNGSLAESEV 269

Qy 240 IRSDFNTNAKTIIVQKESVEINCTRPNNNG-----AGDIRQACHN 281

Db 270 IRSDFNTNAKTIIVQKESVEINCTRPNNNG-----AGDIRQACHN 329

Qy 282 ISRAKMDTLKQIYIKRBEQENKTYIFNHSGGDEPIVHSHFNGGEFFYCNSTOLFNS 341

Db 330 ISGEKNNNTLKQIYIKRBEQENKTYIFNHSGGDEPIVHSHFNGGEFFYCNSTOLFNS 389

Qy 342 TANNTEGSNNTEGNTITLPCRIKQIINMOEYKAMYPPIRGQIRCSSNTIGLLTRD 401

Db 390 TW-NNTIGPNNNG-TITLPCRIKQIINMOEYKAMYPPIRGQIRCSSNTIGLLTRD 447

Qy 402 GGIN-ENGTEIFRPGGDMRDNWSELYKKVVKIEPLGVAPTKCRVYQREKRAVIG 460

Db 448 GKEISNTTEIFRPGGDMRDNWSELYKKVVKIEPLGVAPTKCRVYQREKRAVIG 507

Qy 461 AVFLGFIAGAAGSTMGASMTLTVQARLLLSGIYQOONNLRAIEAOQRMLOLTWGIKOL 520

Db 508 AMFLGFIAGAAGSTMGASMTLTVQARLLLSGIYQOONNLRAIEAOQRMLOLTWGIKOL 567

Qy 521 QARVLAVERYLKRDQQLIGWCGSGKLICTTAVPNNASMSKSLDRINWNTWMEWEIED 580

Db 568 QARVLAVERYLKRDQQLIGWCGSGKLICTTAVPNNASMSKSLDRINWNTWMEWEIED 627

Qy 581 NYTSEIYTLIEESONQOEKNEQELLELDKNASLWNMFDTIKMLWY 625

Db 628 NYTSEIYTLIEESONQOEKNEQELLELDKNASLWNMFDTIKMLWY 672

RESULT 8

AAB61506

AAB61506 standard; Protein; 579 AA.

AC AAB61506;

DT 05-APR-2001 (first entry)

XX

DE HIV-1 deltaV12+ S08 gp140 glycoprotein.

XX

KM gp120; anti-HIV; gene therapy; vaccine; viral envelope protein; gp140;

KM gp140.

XX

XX Human immunodeficiency virus type 1.

OS W0200100648-A1.

XX

PD 04-JAN-2001.

XX

PF 23-JUN-2000; 2000WO-US17267.

XX

PR 25-JUN-1999; 99US-0340992.

XX

PA (PROG-) PROGENICS PHARM INC.

PA (AARO-) AARON DIAMOND AIDS RES CENT.

PI Binley JM, Schuelke N, Olson WC, Maddon PJ, Moore JF;

XX

XX WPI; 2001-122993/13.

DR N-PSDB; AAF28582.

XX

PT New viral envelope proteins, useful for producing vaccines to treat

PT human immunodeficiency virus-1 infections, comprises amino acid
 PT sequence mutations such that viral transmembrane-surface protein
 PT complex is more stable
 XX
 XX
 PS Disclosure; Fig 14; 109pp; English.
 XX
 CC The present invention relates to a viral envelope protein. The viral
 CC envelope protein comprises a viral surface protein (e.g. glycoprotein
 CC gp120) and a corresponding viral transmembrane protein (e.g. gp41), in
 CC which the viral envelope protein contains one or more amino acid sequence
 CC mutations that enhance the stability of the complex formed between the
 CC viral surface and transmembrane proteins. The viral envelope protein can
 CC be used in the treatment of viral infection e.g. HIV-1 infection. The
 CC present sequence is HIV-1 deltaV12* SOS gp140, which was used in the
 CC present invention.
 XX
 XX Sequence 579 AA;
 SO
 Query Match 87.2%; Score 2923; DB 22; Length 579;
 Best Local Similarity 86.0%; Pred. No. 1.3e-163;
 Matches 553; Conservative 0; Mismatches 8; Indels 82; Gaps 2;
 QY 1 VKRLWTVYVYGVWKEATTTFLCASDAKAYTEVHNWATAACVPTDNPQEVLENT 60
 DB 1 VKRLWTVYVYGVWKEATTTFLCASDAKAYTEVHNWATAACVPTDNPQEVLENT 60
 QY 61 EHFNMKNMVEOMODIISLMDOSLKPCKLTPLCVTLNCKVNAITNTNSEGMTERG 120
 DB 61 EHFNMKNMVEOMODIISLMDOSLKPCKLTPLCGA----- 98
 QY 121 EIKNCSFNITTSIRDEVOKEYALFYKLDVYXIDNNNTSYRLISCDSVTITQACPKISFEP 180
 DB 99 -----GCDTSVITQACPKISFEP 116
 QY 181 IPIHYCAPAGFALIKCNDKTFNGKPCCKNVSTQCTHGIRPVVSTOLLNGLAESEVYI 240
 DB 117 IPIHYCAPAGFALIKCNDKTFNGKPCCKNVSTQCTHGIRPVVSTOLLNGLAESEVYI 176
 QY 241 RSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRQAHQNI 282
 DB 177 RSDNFTNNAKTIIVOLKESVEINCTRPNNNTRKSHIGGKAPFTTGELIGDIRQAHQNI 236
 QY 283 SRAKNNDLKOIIVIKLREQFENKTIIVFNHSSGGDEPIVHNSFCGGEFFYCNSSTOLFENST 342
 DB 237 SRAKNNDLKOIIVIKLREQFENKTIIVFNHSSGGDEPIVHNSFCGGEFFYCNSSTOLFENST 296
 QY 343 WNNNTGSSNNTBGNITTLPCRIOIINMOQEVKAMYPPIRQIRCCSNTTGLLITRDG 402
 DB 297 WNNNTGSSNNTBGNITTLPCRIOIINMOQEVKAMYPPIRQIRCCSNTTGLLITRDG 356
 QY 403 GINENGTEIFRPGGGMNRNMRSELYKYVYVKEIPIGVAPTYCKRRVVOREKAVGIGAV 462
 DB 357 GINENGTEIFRPGGGMNRNMRSELYKYVYVKEIPIGVAPTYCKRRVVOREKAVGIGAV 416
 QY 463 FLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOQNNLLRAIEAQRMQLTVWIGIQLOLA 522
 DB 417 FLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOQNNLLRAIEAQRMQLTVWIGIQLOLA 476
 QY 523 RFLAVERVYIGDOOLGIGWCGSKLICTTAVPNNASNSKSLDIRINNNTMWEWEREIDNY 582
 DB 477 RFLAVERVYIGDOOLGIGWCGSKLICTTAVPNNASNSKSLDIRINNNTMWEWEREIDNY 536
 QY 583 TSEIYTLIEESONQOEKNEOBLELDKMASLMMWPIITKMLY 625
 DB 537 TSEIYTLIEESONQOEKNEOBLELDKMASLMMWPIITWLMY 579

RESULT 9
 ID ABU66565 standard; Protein; 842 AA.
 XX AC ABU66565;
 XX

DT 22-MAY-2003 (first entry)
 XX
 XX Human immunodeficiency virus (HIV) envelope (env) protein #1.
 DE
 XX
 XX Immunogenic HIV polypeptide; human immunodeficiency virus; HIV;
 KW vaccine; gene therapy; packaging cell line; humoral immune response;
 KW cellular immune response; gene delivery vector; DNA immunisation;
 KW envelope protein; env.
 OS Human immunodeficiency virus.
 XX
 XX WO2003004657-A1.
 PN
 XX 16-JAN-2003.
 PD
 XX 05-JUL-2002; 2002WO-US21421.
 PF
 XX 05-JUL-2001; 2001US-303192P.
 PR 31-AUG-2001; 2001US-316860P.
 PR 16-JAN-2002; 2002US-349728P.
 PR 16-JAN-2002; 2002US-349793P.
 PR 16-JAN-2002; 2002US-349871P.
 XX
 XX (CHIR) CHIRON CORP.
 PA
 XX
 PI Zur Megele J, Barnett SW, Lian Y;
 DR WPI; 2003-221602/21.
 XX
 PT New synthetic polynucleotides encoding antigenic HIV type B and/or type
 PT C polypeptides, useful as immunogenic compositions or vaccines for
 PT generating humoral or cellular immune responses against HIV in a
 PT subject, especially humans -
 XX
 XX Disclosure; Fig 2A-C; 262pp; English.
 PS
 XX The invention describes a synthetic polynucleotide encoding 2 or more
 CC immunogenic HIV polypeptides, where at least 2 of the polypeptides are
 CC derived from different HIV subtypes. The polynucleotide is useful for
 CC immunisation, generation of packaging cell lines, or production of HIV
 CC polypeptides. The polynucleotide and its encoded proteins are useful as
 CC immunogenic compositions or vaccines for generating humoral or cellular
 CC immune responses against HIV in a subject, or for inducing neutralising
 CC antibodies against HIV. The gene delivery vector comprising the
 CC polynucleotide is also useful for DNA immunisation of, or for
 CC generating an immune response (e.g. a humoral or cellular immune
 CC response) in, a subject such as a mammal, particularly a human. This
 CC is the amino acid sequence of a human immunodeficiency virus (HIV)
 CC envelope (env) protein.
 XX
 XX Sequence 842 AA;
 SO
 Query Match 87.2%; Score 2923; DB 24; Length 842;
 Best Local Similarity 85.9%; Pred. No. 1.8e-163;
 Matches 554; Conservative 28; Mismatches 41; Indels 22; Gaps 5;
 QY 1 VKRLWTVYVYGVWKEATTTFLCASDAKAYTEVHNWATAACVPTDNPQEVLENT 60
 DB 25 VKRLWTVYVYGVWKEATTTFLCASDAKAYTEVHNWATAACVPTDNPQEVLENT 84
 QY 61 EHFNMKNMVEOMODIISLMDOSLKPCKLTPLCVTLNCKVNAITNTNSEGMTERG 119
 DB 85 EHFNMKNMVEOMODIISLMDOSLKPCKLTPLCVTLNCKVNAITNTNSEGMTERG 144
 QY 120 GEIKNCSFNITTSIRDEVOKEYALFYKLDVYXIDNNNTSYRLISCDSVTITQACPKISFE 179
 DB 145 GEIKNCSFNITTSIRDEVOKEYALFYKLDVYXIDNNNTSYRLISCDSVTITQACPKISFE 204
 QY 180 PIPHYCAPAGFALIKCNDKTFNGKPCCKNVSTQCTHGIRPVVSTOLLNGLAESEVYI 239
 DB 205 PIPHYCAPAGFALIKCNDKTFNGKPCCKNVSTQCTHGIRPVVSTOLLNGLAESEVYI 264
 QY 240 IRSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRQAHQNI 281


```

Db 265 IRSEFTDNAKTIIVOLKESVEINCTRPNNTRKSIITIGPGRAFYATGDIIGDIRQAHCN 324
Qy 282 ISRAKMDTLKQIYIKLREOFENKTIYFNHSSGGDEIWMHSFNGGSEFFFCNSTOLFNS 341
Db 325 ISGERKMNNTLKQIYIKLQAOFGNKTIYFKOSSGGDEIWMHSFNGGSEFFFCNSTOLFNS 384
Qy 342 TWNNTEGSNNTGNTITLPCRIKQIIMMOEVEKAMAPPIRGQIRCSSNITGLLTRD 401
Db 385 TW-NNTIGPNTNG-TITLPCRIKQIIMMOEVEKAMAPPIRGQIRCSSNITGLLTRD 442
Qy 402 GGIN-ENGTEIFRPGGDMRDNRSELYKRVKIEPLGVAPTCKRRVQREKRAVIG 460
Db 443 GGEKISNTTEIFRPGGDMRDNRSELYKRVKIEPLGVAPTCKRRVQREKRAVIG 502
Qy 461 AVFLGFLGAAGSTGASMTLTVQARLLSGIVQOONNLRAIEAQOHLQLTWGIKQL 520
Db 503 AMFLGFLGAAGSTGASMTLTVQARLLSGIVQOONNLRAIEAQOHLQLTWGIKQL 562
Qy 521 QARVLAVERYLKQDQQLGIWCGSGKLICTTAVPNNASMSKSLDRINNMTWMEEREID 580
Db 563 QARVLAVERYLKQDQQLGIWCGSGKLICTTAVPNNASMSKSLDRINNMTWMEEREID 622
Qy 581 NYTSEIYTLIESONQOEKNEQELLEDKMASLNNWPDITKMLWY 625
Db 623 NYTMLIYTLIESONQOEKNEQELLEDKMASLNNWPDITKMLWY 667

```

```

RESULT 10
AB06211
ID ABB06211 standard; Protein; 842 AA.
XX
AC ABB06211;
XX
DT 15-MAY-2002 (first entry)
XX
DE HIV Env isolate SF162 amino acid sequence.
XX
KW Human immunodeficiency virus type C; antigenic HIV type C protein;
KW immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW immunostimulant; gene therapy.
XX
OS Human immunodeficiency virus type C.
OS Synthetic.
XX
PN WO200204493-A2.
XX
17-JAN-2002.
XX
05-JUL-2001; 2001WO-US21241.
XX
PR 05-JUL-2000; 2000US-0610313.
XX
PA (CHIR ) CHIRON CORP.
PA (UYST-) UNIV STELLENBOSCH.
XX
PI Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
XX
WPI; 2002-154920/20.
XX
PT New polynucleotides encoding antigenic HIV Type C polypeptides, useful
XX in applications including DNA immunization or generation of packaging
XX cell lines, particularly in gene therapy
XX
PS Disclosure; Fig 105; 233pp; English.
XX
CC The present invention describes expression cassettes comprising a
XX polynucleotide sequence encoding a polypeptide comprising immunoge-
XX nic HIV type C polypeptides. The expression cassettes comprise any of the
XX HIV type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or
XX Nef (1). (1) have immunostimulant activity and can be used in gene
XX therapy. The HIV type C polynucleotides are useful in applications
XX including DNA immunisation, generation of packaging cell lines, and

```

CC production of HIV Type C proteins. The polynucleotides are particularly
 CC useful in gene therapy and DNA immunisation applications. ABL39942 to
 CC ABI40054 and ABB06215 represent sequences used in the
 CC exemplification of the present invention.

XX Sequence 842 AA;

Query Match 87.1%; Score 2918; DB 23; Length 842;
 Best Local Similarity 85.7%; Pred. No. 3,4e-163;
 Matches 553; Conservative 28; Mismatches 42; Indels 22; Gaps 5;

```

Qy 1 VEKLMVTYVYGVVWPKKATITLFCASDAKADTEVHNWMAHACVPTDPNDEVLNVT 60
Db 25 VEKLMVTYVYGVVWPKKATITLFCASDAKADTEVHNWMAHACVPTDPNDEVLNVT 84
Qy 61 EHFMMKNNWVEQNEQDIIISLMDSLKPCVLTPLCYLHCTNKNATKSSWKEMDR 119
Db 85 ENFMWKNNNWVEQNEQDIIISLMDSLKPCVLTPLCYLHCTNKNATKSSWKEMDR 144
Qy 120 GEIKNCFNITTSIRDEVOKEVALFYKLDVYXIDNNNTSYRLISCDTSVITQACPKISFE 179
Db 145 GEIKNCFKVTTSIRKMKQEXALFYKLDVPIINDNTSYRLINCNTSVITQACPKISFE 204
Qy 180 PIPHYCAPAGFALIKCNDKTFNGKXPCKNVSTYQCTHGIRPVVSTOLLNGSLAEEVY 239
Db 205 PIPHYCAPAGFALIKCNDKTFNGKXPCKNVSTYQCTHGIRPVVSTOLLNGSLAEEVY 264
Qy 240 IRSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRQAHCN 281
Db 265 IRSEFTDNAKTIIVOLKESVEINCTRPNNTRKSIITIGPGRAFYATGDIIGDIRQAHCN 324
Qy 282 ISRAKMDTLKQIYIKLREOFENKTIYFNHSSGGDEIWMHSFNGGSEFFFCNSTOLFNS 341
Db 325 ISGERKMNNTLKQIYIKLQAOFGNKTIYFKOSSGGDEIWMHSFNGGSEFFFCNSTOLFNS 384
Qy 342 TWNNTEGSNNTGNTITLPCRIKQIIMMOEVEKAMAPPIRGQIRCSSNITGLLTRD 401
Db 385 TW-NNTIGPNTNG-TITLPCRIKQIIMMOEVEKAMAPPIRGQIRCSSNITGLLTRD 442
Qy 402 GGIN-ENGTEIFRPGGDMRDNRSELYKRVKIEPLGVAPTCKRRVQREKRAVIG 460
Db 443 GGEKISNTTEIFRPGGDMRDNRSELYKRVKIEPLGVAPTCKRRVQREKRAVIG 502
Qy 461 AVFLGFLGAAGSTGASMTLTVQARLLSGIVQOONNLRAIEAQOHLQLTWGIKQL 520
Db 503 AMFLGFLGAAGSTGASMTLTVQARLLSGIVQOONNLRAIEAQOHLQLTWGIKQL 562
Qy 521 QARVLAVERYLKQDQQLGIWCGSGKLICTTAVPNNASMSKSLDRINNMTWMEEREID 580
Db 563 QARVLAVERYLKQDQQLGIWCGSGKLICTTAVPNNASMSKSLDRINNMTWMEEREID 622
Qy 581 NYTSEIYTLIESONQOEKNEQELLEDKMASLNNWPDITKMLWY 625
Db 623 NYTMLIYTLIESONQOEKNEQELLEDKMASLNNWPDITKMLWY 667

```

```

RESULT 11
AAB60422
ID AAB60422 standard; Protein; 868 AA.
XX
AC AAB60422;
XX
DT 25-MAR-2003 (updated)
DT 20-AUG-1991 (first entry)
XX
DE Sequence of LAV virus env protein.
XX
KW AIDS vaccine; diagnosis; immunoassay; HIV; HTLV-III.
XX
OS Lymphadenopathy virus.
XX
PN WO8602383-A.
XX

```


PD 24-APR-1986.
 XX
 PF 18-OCT-1985; 85WO-EP00548.
 XX
 PR 18-OCT-1984; 84GB-0016013.
 PR 16-NOV-1984; 84GB-0029099.
 PR 21-JAN-1985; 85GB-0001473.
 XX
 PA (INSP) INST PASTEUR.
 PA (CNRS) CENT NAT RECH SCIENTIF.
 XX
 PI Montagnier L, Krust B, Charnaret F, Chermann JC, Barresinou F,
 PI Alison M, Sonigo P;
 XX
 DR WPI: 1986-119166/18.
 DR N-PSDB; AAN60365.
 XX
 PT Purified glyco:protein and peptide(s) - are recognised by sera contg.
 PT antibodies against lymphadenopathy virus and useful in detecting
 XX
 AIDS antibodies or in vaccines
 XX
 Disclosure: Fig 4; 75pp; English.
 XX
 CC The inventors claim a polypeptide which is recognised by sera of
 CC human origin contg. antibodies against the virus of
 CC lymphadenopathies (LAV) or acquired immune deficiency syndrome
 CC (AIDS). Also claimed are various peptides corresp. to the AA
 CC sequences deducible from proteins encoded by LAV DNA, defined by
 CC specific residues (e.g. 12-32, 37-46, 49-79, 88-153) in accordance
 CC with a formula given in the specification.
 CC (Updated on 25-MAR-2003 to correct PR field.)
 CC (Updated on 25-MAR-2003 to correct PA field.)
 CC (Updated on 25-MAR-2003 to correct PI field.)
 XX
 SQ Sequence 868 AA;
 Query Match 86.5%; Score 2898.5; DB 7; Length 868;
 Best Local Similarity 84.1%; Pred. No. 4,9e-162;
 Matches 551; Conservative 32; Mismatches 41; Indels 31; Gaps 7;
 QY 2 EKLNTVTVYGVVWKEATTTTFCASDADAKYDTEVNNWATACVPTDPNPQEVLENTVE 61
 DB 39 EKLNTVTVYGVVWKEATTTTFCASDADAKYDTEVNNWATACVPTDPNPQEVLENTVE 98
 QY 62 HNNMMKNWNEVMOEDIIISLMDOSLKPCKLTPLCYTLNCKV-NATN-----TDSSE- 115
 DB 99 NNNMMKNWNEVMOEDIIISLMDOSLKPCKLTPLCYTLNCKV-NATN-----TDSSE- 158
 DB 116 -TMRGEIKNCSFNITTSIRDEVOKEYALFYKLDVYXIDNNNTSYRLISCDTSVITQACP 174
 DB 159 MMRMGEIKNCSFNITTSIRDEVOKEYALFYKLDVYXIDNNNTSYRLISCDTSVITQACP 218
 QY 175 KISFEPPIPHYCAPGFAILKCNKDTFNGKXPCKNVSTYQCHGIRPVVSTOLLNGSLA 234
 DB 219 KVSFEPPIPHYCAPGFAILKCNKDTFNGKXPCKNVSTYQCHGIRPVVSTOLLNGSLA 278
 QY 235 EEEVIRSDNFTNNKTIIVOLKESYEINCTRPNNN-----GAGDI 275
 DB 279 EEEVIRSDNFTNNKTIIVOLKESYEINCTRPNNN-----GAGDI 338
 QY 276 ROAHNCISRAKNDTLKOIVIKLEQF-ENKTIIVENHSGGDEIYVHSGGGEFFYCN 334
 DB 339 ROAHNCISRAKNDTLKOIVIKLEQF-ENKTIIVENHSGGDEIYVHSGGGEFFYCN 398
 QY 335 STOLFNSNTNN--NTEGSSNTEG-NTITLPCRKIQIIMMOEYVKAMYPPIRQIRCS 390
 DB 339 STOLFNSNTNN--NTEGSSNTEG-NTITLPCRKIQIIMMOEYVKAMYPPIRQIRCS 458
 QY 331 SNITGLLLTRDGGIENGEIFRPGGDMRDWRSLYKYVKIPELGAATPKCRRV 450
 DB 459 SNITGLLLTRDGGIENGEIFRPGGDMRDWRSLYKYVKIPELGAATPKCRRV 518
 QY 451 QREKRAVIGAVFLGFLGAGSTMGASMTLTVQARLLISGLVQOONNLRAIEAQORWL 510

DB 519 QREKRAVIGAVFLGFLGAGSTMGASMTLTVQARLLISGLVQOONNLRAIEAQORWL 578
 QY 511 QLTWGIQLOARVAVERYLEDQOLLGWSSGKLTCTTAPNPNASWNSKSLDRIMNM 570
 DB 579 QLTWGIQLOARVAVERYLEDQOLLGWSSGKLTCTTAPNPNASWNSKSLDRIMNM 638
 QY 571 TMMEREIDNTSYLTLIESONQOEKNEBELDPMKASLMMWFDITKWLAMY 625
 DB 639 TMMEREIDNTSYLTLIESONQOEKNEBELDPMKASLMMWFDITKWLAMY 693
 RESULT 12
 AA019389
 ID AA019389 standard; Protein; 868 AA.
 AC AA019389;
 XX
 DT 10-DEC-2002 (first entry)
 XX
 DE Lymphadenopathy-associated virus env protein.
 XX
 KW Lymphadenopathy-associated virus; LAV; HIV; human immunodeficiency virus;
 KW Lymphadenopathy; AIDS; gag; pol; orf Q; env; orf F; anti-HIV; virucide;
 XX
 OS Lymphadenopathy-associated virus.
 PN US6440657-B1.
 PD 27-AUG-2002.
 XX
 PF 06-JAN-2000; 2000US-0478492.
 XX
 PR 16-NOV-1984; 84GB-0029099.
 PR 18-OCT-1985; 85CA-0493377.
 PR 03-NOV-1992; 92US-0970954.
 PR 17-OCT-1986; 86US-0920119.
 PR 01-AUG-1989; 88US-0390499.
 PR 06-DEC-1990; 90US-0622278.
 PR 20-AUG-1991; 91US-0747506.
 PR 01-OCT-1993; 93US-0130565.
 PR 06-JUN-1995; 95US-0466920.
 PR 05-DEC-1993; 83US-0558109.
 PR 28-FEB-1985; 85US-0706562.
 PR 30-AUG-1985; 85US-0771230.
 PR 30-AUG-1985; 85US-0771247.
 PR 30-AUG-1985; 85US-0771248.
 XX
 PA (INSP) INST PASTEUR.
 PA (CNRS) CENT NAT RECH SCI.
 XX
 PI Montagnier L, Krust B, Charnaret S, Clavel F, Chermann J;
 PI Barre-Sinoussi F, Alison M, Sonigo P, Cole S, Danos O;
 PI Wain-Hobson S;
 XX
 DR WPI: 2002-711525/77.
 DR N-PSDB; AAL49920.
 XX
 PT Novel chemically synthesized Human Immunodeficiency Virus envelope
 PT protein, useful in vitro diagnostic method for the detection of the
 PT presence or absence of antibodies that bind to antigens of HIV-1
 PT retrovirus -
 XX
 PS Disclosure: Fig 3; 31pp; English.
 XX
 CC The present invention relates to a chemically synthesised env peptide of
 CC Human Immunodeficiency Virus (HIV) of less than 150 amino acid residues.
 CC Such a peptide is useful in an in vitro diagnostic method for the
 CC detection of the presence or absence of antibodies that bind to antigens
 CC of HIV-1 retrovirus, by contacting the peptide with a biological fluid
 CC for a time and under conditions sufficient for the peptide and antibody
 CC in the biological fluid to form a peptide-antibody complex, detecting the

formation of the peptide-antibody complex by comparing the formation of peptide-antibody complex formation with a control sample, where the formation of peptide-antibody complex is correlated with the presence of antibodies that bind to antigens of HIV-1 retrovirus in the biological sample. It can also be used for preparation of vaccine compositions against AIDS or related syndromes. The present sequence is the CC lymphadenopathy-associated virus LAV env protein. The LAV virus was isolated from a patient with AIDS.

XX Sequence 868 AA;

Query Match 86.4%; Score 2895.5; DB 23; Length 868;
Best Local Similarity 84.1%; Pred. No. 7.4e-162;
Matches 551; Conservative 31; Mismatches 42; Indels 31; Gaps 7;

```

QY 2 EKLWVTVYGVVWKEATTTLPFCASDAKAYDTEVHNVWATHACVPTDPNPEVLENTTE 61
DB 39 EKLWVTVYGVVWKEATTTLPFCASDAKAYDTEVHNVWATHACVPTDPNPEVLENTTE 98
62 HNNMKNMVEOMEDIIISLDOSLPCVKLTPLCVTLNCKDV-NATNT---TNDSEG- 115
99 NFNMMKNDMVEOMEDIIISLDOSLPCVKLTPLCVSLKCTDLGNATNTSSNTSSGE 158
QY 116 -TMRGEIKNCSFNITTSIRDEVQKEYALFYKLDVXIDNNNTSYRLISCDTSVITQACP 174
DB 159 MMEKEIKNCSFNITTSIRDEVQKEYAFYKLDIIPIDNDTSTVLTSCNTSVITQACP 218
QY 175 KISEPPIPIHCAAGRAIILKNDKTFNGKPKCKNVSTXOCTHGIRPVVSTOLLNGLA 234
DB 219 KISEPPIPIHCAAGRAIILKNDKTFNGKPKCKNVSTXOCTHGIRPVVSTOLLNGLA 278
QY 235 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNN-----GAGDI 275
DB 279 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNN-----GAGDI 338
QY 276 ROACHNISRAKMDTLKOIVIKLRQF-ENKTIVENHSSGGDPEIWMHSFNGGGEFFYCN 334
DB 339 ROACHNISRAKMDTLKOIVIKLRQF-ENKTIVENHSSGGDPEIWMHSFNGGGEFFYCN 398
QY 335 STOLFNSITWNN---NTBGSNTEG-NTITLPCRIOIINMWOEVKAMYPAPIRGOIRCS 390
DB 399 STOLFNSITWNN---NTBGSNTEG-NTITLPCRIOIINMWOEVKAMYPAPIRGOIRCS 458
QY 391 SNITGLLTRDGINENGTEIFRPGGDMRDNMSELYKYKVKIEPLGVAPTKCKRRV 450
DB 459 SNITGLLTRDGINENGTEIFRPGGDMRDNMSELYKYKVKIEPLGVAPTKCKRRV 518
QY 451 QREKAVGIGAVFLGFLGAGSTWGAASMTLTVQARLLLSGIVOOONNLLRAIEAQORML 510
DB 519 QREKAVGIGAVFLGFLGAGSTWGAASMTLTVQARLLLSGIVOOONNLLRAIEAQORML 578
QY 511 QLTWGIQOLQARUAVERYLGDQOLLGIWCGSGULICTAVPNNASNSKSLDINNM 570
DB 579 QLTWGIQOLQARUAVERYLGDQOLLGIWCGSGULICTAVPNNASNSKSLDINNM 638
QY 571 TWMEVEREIDNVTSEIYTLIESONQOEKNEBELDKMASLWMPDITKMLY 625
DB 639 TWMEVEREIDNVTSEIYTLIESONQOEKNEBELDKMASLWMPDITKMLY 693

```

RESULT 13
AAP60063
ID AAP60063 standard; Protein; 868 AA.

XX AAP60063;
AC 25-MAR-2003 (updated)
XX 31-OCT-2002 (updated)
DT 11-JUN-1991 (first entry)
XX HIV virus env gene product.
DE
XX HIV virus; vaccine; AIDS.
KW

XX OS Human immunodeficiency virus.

XX FN EP201716-A.

XX PD 20-NOV-1986.

XX PF 07-APR-1986; 86EP-0104741.

XX PR 07-NOV-1985; 85US-0783229.

XX PR 08-APR-1985; 85US-0721237.

XX PR 29-APR-1985; 85US-0728052.

XX PR 10-JUL-1985; 85US-0753769.

XX (GENE-) GENETIC SYSTEMS CORP.

XX (BRIM) BRISTOL MYERS SQUIBB CO.

XX Watanabe SM, Cosand W, Mardale S, Travis BW;

XX N-PSDB; AAN60077.

XX NPI; 1986-306545/47.

XX New DNA sequence comprising part of env. region of LAV genome -

XX PT and corresp. transformational protein prod., useful in vaccine

XX against AIDS.

XX Disclosure; Fig 8; 27pp; English.

XX The protein is immunologically reactive with antibodies against

XX HIV virus, and can be used in a vaccine composition for

XX prevention against AIDS.

XX (Updated on 31-OCT-2002 to add missing OS field.)

XX (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 868 AA;

Query Match 86.4%; Score 2894.5; DB 7; Length 868;
Best Local Similarity 84.1%; Pred. No. 8.4e-162;
Matches 551; Conservative 31; Mismatches 42; Indels 31; Gaps 7;

```

QY 2 EKLWVTVYGVVWKEATTTLPFCASDAKAYDTEVHNVWATHACVPTDPNPEVLENTTE 61
DB 39 EKLWVTVYGVVWKEATTTLPFCASDAKAYDTEVHNVWATHACVPTDPNPEVLENTTE 98
62 HNNMKNMVEOMEDIIISLDOSLPCVKLTPLCVTLNCKDV-NATNT---TNDSEG- 115
99 NFNMMKNDMVEOMEDIIISLDOSLPCVKLTPLCVSLKCTDLGNATNTSSNTSSGE 158
QY 116 -TMRGEIKNCSFNITTSIRDEVQKEYALFYKLDVXIDNNNTSYRLISCDTSVITQACP 174
DB 159 MMEKEIKNCSFNITTSIRDEVQKEYAFYKLDIIPIDNDTSTVLTSCNTSVITQACP 218
QY 175 KISEPPIPIHCAAGRAIILKNDKTFNGKPKCKNVSTXOCTHGIRPVVSTOLLNGLA 234
DB 219 KISEPPIPIHCAAGRAIILKNDKTFNGKPKCKNVSTXOCTHGIRPVVSTOLLNGLA 278
QY 235 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNN-----GAGDI 275
DB 279 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNN-----GAGDI 338
QY 276 ROACHNISRAKMDTLKOIVIKLRQF-ENKTIVENHSSGGDPEIWMHSFNGGGEFFYCN 334
DB 339 ROACHNISRAKMDTLKOIVIKLRQF-ENKTIVENHSSGGDPEIWMHSFNGGGEFFYCN 398
QY 335 STOLFNSITWNN---NTBGSNTEG-NTITLPCRIOIINMWOEVKAMYPAPIRGOIRCS 390
DB 399 STOLFNSITWNN---NTBGSNTEG-NTITLPCRIOIINMWOEVKAMYPAPIRGOIRCS 458
QY 391 SNITGLLTRDGINENGTEIFRPGGDMRDNMSELYKYKVKIEPLGVAPTKCKRRV 450
DB 459 SNITGLLTRDGINENGTEIFRPGGDMRDNMSELYKYKVKIEPLGVAPTKCKRRV 518
QY 451 QREKAVGIGAVFLGFLGAGSTWGAASMTLTVQARLLLSGIVOOONNLLRAIEAQORML 510

```

```

Db      519 OREKAVGALFLFELGAGSTMGARSMTLTVQARQLLSGIVQOONLRAIEAQHLL 578
Qy      511 QLTWGIKQLOARVLAVERYLGDQOLGIGWCSGKLICTTAVPMNASWSNKSLEDRIMNMN 570
Db      579 QLTWGIKQLOARVLAVERYLGDQOLGIGWCSGKLICTTAVPMNASWSNKSLEQIMNMN 638
Qy      571 TWMEREIDNNTSEIYTLIEESONQOEKNEQELBELDKWASLMMWFDTIKWLMY 625
Db      639 TWMERDREINNTSLIHSLEESONQOEKNEQELBELDKWASLMMWFNTTNLMY 693

RESULT 14
AAB85999 standard; Protein; 856 AA.
ID      AAB85999
AC      AAB85999;
XX      30-NOV-2001 (first entry)
DT
XX      Amino acid sequence of HIV-1 isolate BH10 envelope polypeptide.
XX      HIV-1; gp120; BH10; vaccine; immunization; envelope polypeptide.
XX      Human immunodeficiency virus type 1.
XX      Key Location/Qualifiers
XX      FT Peptide 79..184
XX      FT Peptide /note= "specifically claimed fragment"
XX      FT Peptide 326..400
XX      FT Peptide /note= "specifically claimed fragment"
XX      US6268484-B1.
XX      31-JUL-2001.
XX      30-JUL-1998; 98US-0124900.
XX      07-JUN-1995; 95US-0478536.
XX      19-APR-1995; 95WO-EP01481.
XX      (POLY-) POLY(MN SCI IMMOBIOLOGISCHE FORSCHUNG.
XX      Katinger H, Buchacher A, Ernst W, Ballaun C, Purtscher M;
XX      Trkola A, Predl R, Schmatz C, Klima A, Steindl F, Muster T;
XX      WPI; 2001-556601/62.
XX      N-PSDB; AAH76385.
XX      New anti-idiotypic antibodies consisting of one or both amino acid
XX      sequences corresponding to amino acid positions 79-84 or 326-400 of the
XX      processed gp120 of HIV-isolate BH10, useful as vaccine against HIV-1
XX      infections.
XX      Claim 1; Columns 33-38; 27pp; English.
XX      The invention relates to a peptide fragment comprising of amino acid
XX      sequences corresponding to sequences within the processed gp120 of HIV-1
XX      isolate BH10 (Genbank accession M15654). The peptides are useful in the
XX      detection, prevention and treatment of HIV-1 infections, and in AIDS
XX      therapy. The antibodies are especially useful as vaccines for active and
XX      passive immunization, or for the detection and/or determination of HIV-1
XX      infected cells and/or HIV-1 viruses. The present sequence represents the
XX      amino acid sequence of a envelope polypeptide of HIV-1 isolate BH10
XX      (Genbank accession M15654).
XX      Sequence 856 AA;
SQ
Query Match 86.2%; Score 2887; DB 22; Length 856;
Best Local Similarity 84.0%; Pred. No. 2,3e-161;
Matches 546; Conservative 34; Mismatches 44; Indels 26; Gaps 5;
Qy      2 EKLMTVYVGVVWKEATTTLFCASDAKAYDTEVHNWVATHACVPTDPNPQEVVLVWTE 61

```

```

Db      32 EKLMTVYVGVVWKEATTTLFCASDAKAYDTEVHNWVATHACVPTDPNPQEVVLVWTE 91
Qy      62 HNNMKNMNMREOMODIISIMDOSLKPCVKLTPLCVTLNCKOVNMTNTNDEG--TMR 119
Db      92 NNNMKNMNMREOMODIISIMDOSLKPCVKLTPLCVTLNCKOVNMTNTNDEG--TMR 151
Qy      120 GEIKNCSFNITTSIRDEVOKEVALFYKLDVVAIDNNNTSRYLISCDTSVITQACPISFE 179
Db      132 GEIKNCSFNITTSIRDEVOKEVALFYKLDVVAIDNNNTSRYLISCDTSVITQACPISFE 211
Qy      180 PIPHYCAPAGFAILIKCKDKTFNGKPCKNVSTXOCTGIRPVSTOLLNLSLAEVY 239
Db      212 PIPHYCAPAGFAILIKCKDKTFNGKPCKNVSTXOCTGIRPVSTOLLNLSLAEVY 271
Qy      240 IRSDFNTNNAKTIYQLAESVEINCTRNPN-----GAGDIRQAMC 280
Db      272 IRSANFTNNAKTIYQLAESVEINCTRNPNNRKSIIRIQEGRAFTYIGKIGNRQAMC 331
Qy      281 NISRAKNMNDTLKQIVIKLREOF-ENKTIYFNHSSGDEPEIVHSPNCGEFPYCNSLOLF 339
Db      332 NISRAKNMNDTLKQIVIKLREOF-ENKTIYFNHSSGDEPEIVHSPNCGEFPYCNSLOLF 391
Qy      340 NSTWNN--NTEGSNNTEG-NTITLPCRIKOIINMGEVKAAMYAPIRGOIRCSNITG 395
Db      392 NSTWNN--NTEGSNNTEG-NTITLPCRIKOIINMGEVKAAMYAPIRGOIRCSNITG 451
Qy      396 LLITRDGGINENGTIFRPGGDMNDNRSELYKRVYKIEPLGVAFTPKCKRRVYQREKR 455
Db      452 LLITRDGGINENGTIFRPGGDMNDNRSELYKRVYKIEPLGVAFTPKCKRRVYQREKR 511
Qy      456 AVGIGAVFLGFLGAGSTMGASMTLTVQARLLSGIVQOONLRAIEAQHLLQLTW 515
Db      512 AVGIGAVFLGFLGAGSTMGASMTLTVQARLLSGIVQOONLRAIEAQHLLQLTW 571
Qy      516 GIKQLOARVLAVERYLGDQOLGIGWCSGKLICTTAVPMNASWSNKSLEDRIMNMN 575
Db      572 GIKQLOARVLAVERYLGDQOLGIGWCSGKLICTTAVPMNASWSNKSLEQIMNMN 631
Qy      576 EREIDNNTSEIYTLIEESONQOEKNEQELBELDKWASLMMWFDTIKWLMY 625
Db      632 EREIDNNTSEIYTLIEESONQOEKNEQELBELDKWASLMMWFNTTNLMY 681

RESULT 15
AAR43869 standard; Protein; 863 AA.
ID      AAR43869
AC      AAR43869;
XX      25-MAR-2003 (updated)
XX      23-DEC-1993 (first entry)
XX      HTLV-III ENV-LOR gene product (Clone BH10).
XX      Polypeptide; antibodies; HTLV; AIDS; vaccine.
XX      Human T-cell lymphotropic virus.
XX      EP552850-A1.
XX      28-JUL-1993.
XX      10-OCT-1985; 93EP-0200929.
XX      10-OCT-1984; 84US-0659339.
XX      23-JAN-1985; 85US-0693866.
XX      10-OCT-1985; 85EP-0307260.
XX      (GENZ ) CENTOCOR INC.
XX      Chang NT, Gallo RC, Wong-staal F;
XX

```

DR WPI: 1993-236543/30.
DR N-PSDB: AAQ45919.

XX Cloning and expression of new HTLV-III DNA - used to obtain
PT polypeptide(s) and antibodies for diagnosis, prevention and
PT treatment of HTLV-III infection, partic. AIDS

XX Disclosure; Figure 3; 31pp; English.

XX A fragment of DNA approximately 200-500 base pairs in length is
CC ligated into a vector and used to transform E.coli. These cells then
CC express a polypeptide which is immunoreactive with HTLV-III-specific
CC antibody. The HTLV-III polypeptides can be used for the production
CC of antibodies, in immunoassays for the detection of HTLV-
CC III-specific antibodies and in vaccines for the prevention of AIDS.
CC The antibodies can also be used to detect HTLV-III polypeptides.
CC (Updated on 25-MAR-2003 to correct PN field.)
CC (Updated on 25-MAR-2003 to correct PR field.)
CC (Updated on 25-MAR-2003 to correct PR field.)

Sequence 863 AA:

Query Match 86.2%; Score 2887; DB 14; Length 863;
Best Local Similarity 84.0%; Pred. No. 2.3e-161;
Matches 546; Conservative 34; Mismatches 44; Indels 26; Gaps 5;

```
QY 2 EKLWTVYVYGVPMWEATTTLFCASDAKAYDTEVNNWATHACVPTDPNPOEVLENTE 61
DB 39 EKLWTVYVYGVPMWEATTTLFCASDAKAYDTEVNNWATHACVPTDPNPOEVLENTE 98
QY 62 HFNWKNMVEOMODIISLWDOSLPCVKLTPLCVTLNCKDVNATNTTNDSEG--TWER 119
DB 99 NFNWKNMVEOMODIISLWDOSLPCVKLTPLCVTLNCKDVNATNTTNDSEG--TWER 158
QY 120 GEIKKCSFNITTSIRDEVQKEYALFYKLDVAXIDNNNTSYRLISCDTSVITQACPXISFE 179
DB 159 GEIKKCSFNITTSIRDEVQKEYALFYKLDVAXIDNNNTSYRLISCDTSVITQACPXISFE 218
QY 180 PIPHYCAPAGPAILKCKDKTFNGKXPKGNVSTXCTHGIRPVYSTOLLNGSLAEEV 239
DB 219 PIPHYCAPAGPAILKCKDKTFNGKXPKGNVSTXCTHGIRPVYSTOLLNGSLAEEV 278
QY 240 IRSDFNNAKTIIVOLKESVEINCTRPNN-----GAGDIRQAH 280
DB 279 IRSDFNNAKTIIVOLKESVEINCTRPNN-----GAGDIRQAH 338
QY 281 NISRAKNDTLKQIVIKLREOF-ENKTIVENHSSGDPPEIVMESHFNCGEFPYCNSTOLF 339
DB 339 NISRAKNDTLKQIVIKLREOF-ENKTIVENHSSGDPPEIVMESHFNCGEFPYCNSTOLF 398
QY 340 NSTWNN--NTEGSNNTGEG-NTITLPCRITKOINNMVGVKAMVAPRIGQIRCSNITG 395
DB 399 NSTWNN--NTEGSNNTGEG-NTITLPCRITKOINNMVGVKAMVAPRIGQIRCSNITG 458
QY 396 LLITFDGGINENGTEIFRPGGDMRDNMRSELYKTVKIEPLGVAPTCKRRVYVQREKR 455
DB 459 LLITFDGGINENGTEIFRPGGDMRDNMRSELYKTVKIEPLGVAPTCKRRVYVQREKR 518
QY 456 AVGIGAVFLGFLGAGSTGMAASMTLTVQARILLSGIVQOONNLLRAIEAQRMILQITW 515
DB 519 AVGIGAVFLGFLGAGSTGMAASMTLTVQARILLSGIVQOONNLLRAIEAQRMILQITW 578
QY 516 GIKOLQARVLAVERLYGDOQLIGTGGCGKLTCTAVPMNASWSKSLDRIINNMNMW 575
DB 579 GIKOLQARVLAVERLYGDOQLIGTGGCGKLTCTAVPMNASWSKSLDRIINNMNMW 638
QY 576 EREIDNYSLEYTLIESONQEKNEOELELDKNASLNMWFDITKMLY 625
DB 639 EREIDNYSLEYTLIESONQEKNEOELELDKNASLNMWFDITKMLY 688
```

Search completed: December 12, 2003, 12:30:07
Job time : 39.2225 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: December 12, 2003, 12:28:04 ; Search time 14.2122 Seconds
(without alignments)
1860.670 Million cell updates/sec

Title: US-10-032-162-17

Perfect score: 3351

Sequence: 1 VEKLWVTVYGVVWKEATT.....ELDKWASLWVFDITKLMY 625

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Minimum number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued Patents AA:*
2: /cgn2_6/prodata/1/aa/5A_COMB.pep:*
3: /cgn2_6/prodata/1/aa/5A_COMB.pep:*
4: /cgn2_6/prodata/1/aa/6A_COMB.pep:*
5: /cgn2_6/prodata/1/aa/6A_COMB.pep:*
6: /cgn2_6/prodata/1/aa/6A_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2968.5	88.6	850	2 US-08-448-603A-28	Sequence 28, App1
2	2968.5	88.6	850	3 US-09-134-075-28	Sequence 28, App1
3	2968.5	88.6	850	4 US-09-492-739-28	Sequence 28, App1
4	2938.5	87.7	855	1 US-08-022-835-6	Sequence 6, App1
5	2938.5	87.7	855	1 US-08-388-809-6	Sequence 6, App1
6	2938.5	87.7	855	2 US-08-647-714-6	Sequence 6, App1
7	2903.5	86.6	861	1 US-07-956-483-14	Sequence 14, App1
8	2903.5	86.6	861	1 US-08-127-499A-14	Sequence 14, App1
9	2901.5	86.6	861	1 US-08-482-847-14	Sequence 14, App1
10	2901.5	86.6	861	3 US-07-956-483-10	Sequence 10, App1
11	2901.5	86.6	861	3 US-08-472-240A-1	Sequence 10, App1
12	2901.5	86.6	861	3 US-08-472-240A-7	Sequence 10, App1
13	2901.5	86.6	861	4 US-08-817-441-103	Sequence 103, App1
14	2893.5	86.2	863	3 US-07-956-483-16	Sequence 16, App1
15	2887	86.2	863	3 US-09-124-800-9	Sequence 9, App1
16	2887	86.2	863	3 US-08-463-210-11	Sequence 11, App1
17	2886	86.1	826	1 US-08-375-510-2	Sequence 2, App1
18	2886	86.1	826	2 US-08-487-657-2	Sequence 2, App1
19	2886	86.1	826	4 US-09-309-572-23	Sequence 23, App1
20	2878	85.9	856	4 US-07-916-098A-2	Sequence 2, App1
21	2876	85.8	857	2 US-08-448-603A-30	Sequence 30, App1
22	2876	85.8	857	3 US-09-134-075-30	Sequence 30, App1
23	2876	85.8	857	4 US-09-492-739-30	Sequence 30, App1
24	2872.5	85.7	839	1 US-08-472-240A-10	Sequence 10, App1
25	2870	85.6	856	1 US-08-022-835-2	Sequence 2, App1
26	2870	85.6	856	1 US-08-388-809-2	Sequence 2, App1
27	2870	85.6	856	2 US-08-647-714-2	Sequence 2, App1

28	2870	85.6	856	3 US-07-956-483-11	Sequence 11, App1
29	2867	85.6	856	4 US-09-337-387-11	Sequence 11, App1
30	2861	85.4	857	1 US-08-022-835-4	Sequence 4, App1
31	2861	85.4	857	1 US-08-388-809-4	Sequence 4, App1
32	2861	85.4	857	2 US-08-647-714-4	Sequence 4, App1
33	2854.5	85.2	880	3 US-08-788-815-7	Sequence 7, App1
34	2854.5	85.2	880	3 US-09-157-963-7	Sequence 7, App1
35	2852	85.1	887	3 US-08-472-240A-5	Sequence 5, App1
36	2852	84.4	855	3 US-07-956-483-15	Sequence 15, App1
37	2822.5	84.2	726	4 US-09-337-387-3	Sequence 3, App1
38	2817	84.1	856	1 US-08-375-100-1	Sequence 1, App1
39	2817	84.1	865	3 US-07-956-483-13	Sequence 13, App1
40	2815.5	84.0	759	4 US-09-337-387-12	Sequence 12, App1
41	2809	83.8	887	3 US-08-472-240A-2	Sequence 2, App1
42	2797	83.5	887	3 US-08-472-240A-6	Sequence 6, App1
43	2788	83.2	887	3 US-08-472-240A-4	Sequence 4, App1
44	2760.5	82.4	665	2 US-08-448-603A-32	Sequence 32, App1
45	2760.5	82.4	665	3 US-09-134-075-32	Sequence 32, App1

ALIGNMENTS

RESULT 1
US-08-448-603A-28
; Sequence 28, Application US/08448603A
; Patent No. 5864027
; GENERAL INFORMATION:
; APPLICANT: Berman, Phillip W.
; APPLICANT: Nakamura, Gerald R.
; TITLE OF INVENTION: HIV Envelope Polypeptides
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSER: McClutchen, Doyle, Brown & Eneersen, LLP
; STREET: 3 Embarcadero Center
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,603A
; FILING DATE: 07-JUN-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/072,833
; FILING DATE: 07-JUN-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Haliday, Emily
; REGISTRATION NUMBER: 38903
; REFERENCE/DOCKET NUMBER: 14918-704
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-393-2000
; TELEFAX: 415-393-2286
; TELEX:
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 850 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-448-603A-28
Query Match 88.6%; Score 2968.5; DB 2; Length 850;
Best Local Similarity 86.9%; Pred. No. 8.8e-248;
Matches 563; Conservative 25; Mismatches 33; Indels 27; Gaps 6;
2 EKLWVTVYGVVWKEATTTTCASDAKAYDIEVHNWATACVPTDPNQEVLNATVE 61
|||||

```

Db      31 EKLWTVYVGVPMKEATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPQEIENYTE 90
Qy      62 HNNMKNMVEQWEDIIISLDQSLKPCVKLTPLCVTLNCKDV-NATNTTNDSEGTMRG 120
Db      91 NFNMMKNMVEQWEDIIISLDQSLKPCVKLTPLCVTLNCKDV-NATNTTNDSEGTMRG 150
Qy      121 EIKKCSFNITTSIRDEVOKEAYLFYKLDVXIDNNNTSYRLISCDTSVITQACPKISFEP 180
Db      151 EIKKCSFNITTSIRDKMKNAYALFYKLDVPIIDNNTSYRLISCDTSVITQACPKISFEP 210
Qy      181 IPIHYCAPAGAILKCDNDKFNKGPCKNVSTXOCTHGIRPVVSTOLLNGLSLAEVEVI 240
Db      211 IPIHYCAPAGAILKCDNDKFNKGPCKNVSTXOCTHGIRPVVSTOLLNGLSLAEVEVI 270
Qy      241 RSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRQAHONI 282
Db      271 RSNFSDNNAKTIIVOLKESVEINCTRPNNNTRRSIHIGPRAFYATGEIIGDIRQAHONI 330
Qy      283 SRAKNDTLKOIIVKLAEPENKTIIVFNHSSGDEPEIVMHSFNGGFEFFYCNSTQLPNST 342
Db      331 SSTKNNTLKOIIVKLAEPENKTIIVFNHSSGDEPEIVMHSFNGGFEFFYCNSTQLPNST 389
Qy      343 WN-----NTEGSNNTEGNTITLPCRIOIIMMVEGKAMAYAPPIRGQIRCSNITGL 397
Db      390 WNTYTYNNTEGSDNT-GRNITLQCRIOIIMMVEGKAMAYAPPIRGQIRCSNITGL 448
Qy      398 LTRDGINENGTETFRPGGDMRDNRSELYKYVVKIEPLGVAPTYKRRVVOREKRAV 457
Db      449 LTRDGG-NNSETEIFRPGGDMRDNRSELYKYVVKIEPLGVAPTYKRRVVOREKRAV 507
Qy      458 GIGAVFLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOONNLRAIEAQOBLQLTWGI 517
Db      508 GIGAVFLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOONNLRAIEAQOBLQLTWGI 567
Qy      518 KOLQARVLAVERYLQDQOLLGIWCSGKLICTTAVPNNASWSNKLDRIMNMNTWEMER 577
Db      568 KOLQARVLAVERYLQDQOLLGIWCSGKLICTTAVPNNASWSNKLDRIMNMNTWEMER 627
Qy      578 EIDNVTSEIYTLIEESQNOQEKNEQELLELDKMASLNMWPDITKWLWY 625
Db      628 EIDNVTSEIYTLIEESQNOQEKNEQELLELDKMASLNMWPDITKWLWY 675

RESULT 2
US-09-134-075-28
; Sequence 28, Application US/09134075
; Patent No. 6042836
; GENERAL INFORMATION:
; APPLICANT: Berman, Phillip W.
; APPLICANT: Nakamura, Gerald R.
; TITLE OF INVENTION: HIV Envelope Polypeptides
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McCutchen, Doyle, Brown & Enersen, LLP
; STREET: 3 Embarcadero Center
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/134,075
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/448,603
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Haliday, Emily

```

```

; REGISTRATION NUMBER: 38903
; REFERENCE/DOCKET NUMBER: 14918-704
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-393-2000
; TELEFAX: 415-393-2286
; TELE:
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 850 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-134-075-28

Query Match      88.6%; Score 2968.5; DB 3; Length 850;
Best Local Similarity 86.9%; Pred. No. 8.8e-248;
Matches 563; Conservative 25; Mismatches 33; Indels 27; Gaps 6;

Qy      2 EKLWTVYVGVPMKEATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPQEIENYTE 61
Db      31 EKLWTVYVGVPMKEATTTLLFCASDAKAYDTEVHNWATHACVPTDPNPQEIENYTE 90
Qy      62 HNNMKNMVEQWEDIIISLDQSLKPCVKLTPLCVTLNCKDV-NATNTTNDSEGTMRG 120
Db      91 NFNMMKNMVEQWEDIIISLDQSLKPCVKLTPLCVTLNCKDV-NATNTTNDSEGTMRG 150
Qy      121 EIKKCSFNITTSIRDEVOKEAYLFYKLDVXIDNNNTSYRLISCDTSVITQACPKISFEP 180
Db      151 EIKKCSFNITTSIRDKMKNAYALFYKLDVPIIDNNTSYRLISCDTSVITQACPKISFEP 210
Qy      181 IPIHYCAPAGAILKCDNDKFNKGPCKNVSTXOCTHGIRPVVSTOLLNGLSLAEVEVI 240
Db      211 IPIHYCAPAGAILKCDNDKFNKGPCKNVSTXOCTHGIRPVVSTOLLNGLSLAEVEVI 270
Qy      241 RSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRQAHONI 282
Db      271 RSNFSDNNAKTIIVOLKESVEINCTRPNNNTRRSIHIGPRAFYATGEIIGDIRQAHONI 330
Qy      283 SRAKNDTLKOIIVKLAEPENKTIIVFNHSSGDEPEIVMHSFNGGFEFFYCNSTQLPNST 342
Db      331 SSTKNNTLKOIIVKLAEPENKTIIVFNHSSGDEPEIVMHSFNGGFEFFYCNSTQLPNST 389
Qy      343 WN-----NTEGSNNTEGNTITLPCRIOIIMMVEGKAMAYAPPIRGQIRCSNITGL 397
Db      390 WNTYTYNNTEGSDNT-GRNITLQCRIOIIMMVEGKAMAYAPPIRGQIRCSNITGL 448
Qy      398 LTRDGINENGTETFRPGGDMRDNRSELYKYVVKIEPLGVAPTYKRRVVOREKRAV 457
Db      449 LTRDGG-NNSETEIFRPGGDMRDNRSELYKYVVKIEPLGVAPTYKRRVVOREKRAV 507
Qy      458 GIGAVFLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOONNLRAIEAQOBLQLTWGI 517
Db      508 GIGAVFLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOONNLRAIEAQOBLQLTWGI 567
Qy      518 KOLQARVLAVERYLQDQOLLGIWCSGKLICTTAVPNNASWSNKLDRIMNMNTWEMER 577
Db      568 KOLQARVLAVERYLQDQOLLGIWCSGKLICTTAVPNNASWSNKLDRIMNMNTWEMER 627
Qy      578 EIDNVTSEIYTLIEESQNOQEKNEQELLELDKMASLNMWPDITKWLWY 625
Db      628 EIDNVTSEIYTLIEESQNOQEKNEQELLELDKMASLNMWPDITKWLWY 675

RESULT 3
US-09-492-739-28
; Sequence 28, Application US/09492739
; Patent No. 6331404
; GENERAL INFORMATION:
; APPLICANT: Berman, Phillip W.
; APPLICANT: Nakamura, Gerald R.
; TITLE OF INVENTION: HIV Envelope Polypeptides
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:

```

ADDRESSEE: McCutchen, Doyle, Brown & Eversen, LLP
STREET: 3 Embarcadero Center
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/492,739
FILING DATE: 27-Jan-2000
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/134,075
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Haliday, Emily
REGISTRATION NUMBER: 38903
REFERENCE/DOCKET NUMBER: 14918-704

TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-393-2000
TELEFAX: 415-393-2286
TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 850 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 28:
US-09-492-739-28

Query Match
Best Local Similarity 88.6%; Score 2968.5; DB 4; Length 850;
Matches 563; Conservative 25; Mismatches 33; Indels 27; Gaps 6;

2 EKLMTVVYGVVPMKEATTLFCASADAKAYDTEVHNWATACVPTDPNPOEVLNVT 61
31 EKLMTVVYGVVPMKEATTLFCASADAKAYDTEVHNWATACVPTDPNPOEVLNVT 90
62 HNNMKNMNVEMOEDIIISLMDQSLKPCVKLTPLCVTLNCKDV-NATNTNDSG 120
91 NNNMKNMNVEMOEDIIISLMDQSLKPCVKLTPLCVTLNCKDV-NATNTNDSG 150
121 EIKNSFNITTSIRDEVOKEVALPYKLDVYXIDNNNTSYRLISCDTSVITQACPKISFEP 180
151 EIKNSFNITTSIRDEVOKEVALPYKLDVYXIDNNNTSYRLISCDTSVITQACPKISFEP 210
181 IPIHICAPAGFAILKNDKTFNGKACKAVSTYQCTHGIRPVYSTOLLNGSLAEVEVY 240
211 IPIHICAPAGFAILKNDKTFNGKACKAVSTYQCTHGIRPVYSTOLLNGSLAEVEVY 270
241 REDNFTNNAKTIIVOLKESEVINCTRPNNNG-----ASDIRAHNCY 282
271 KSAKNSDNKTIIVOLKESEVINCTRPNNNG-----ASDIRAHNCY 330
283 SRAKNDTLKQIVIKLREOPENKTIIVFNHSSGDPRIVHNSFNCGSEFFYCNSTOLFENST 342
331 SSKTNNTLQIVIKLREOPENKTIIVFNHSSGDPRIVHNSFNCGSEFFYCNSTOLFENST 389
343 WY-----NNTGSSNTEGTTITLPCRIKQIIMMGOVGAQAPPIRGQIRCSSNTGL 397
390 WYITVYTNNTGSSNNTI-GNITLQCIKQIIMMGOVGAQAPPIRGQIRCSSNTGL 448
398 LTRDGINENGTEIFRPGGDMRDNRSELYKYKVKIPIPLGAPTKCGRVVOREKRAV 457
449 LTRDGG-ANSETEIFRPGGDMRDNRSELYKYKVKIPIPLGAPTKCGRVVOREKRAV 507
458 GIGAVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOQNNLLRALEAOQRMQLTVMCI 517

508 GIGAVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOQNNLLRALEAOQRMQLTVMCI 567
518 KOLARVLAVERYLGDQOOLLGIMGSGKLICTTAVPANASNSKSLDRIMNTTMMEMER 577
568 KOLARVLAVERYLGDQOOLLGIMGSGKLICTTAVPANASNSKSLDRIMNTTMMEMER 627
578 EIDNNTSEIYTLIBESONQOEKNBOELLELDKMASLWNPDIITKWLWY 625
628 EIDNNTSLIYSLIBESONQOEKNBOELLELDKMASLWNPDIITKWLWY 675

RESULT 4
US-08-022-835-6
Sequence 6, Application US/08022835
Patent No. 5420030

GENERAL INFORMATION:
APPLICANT: Reitz Jr., Marvin S.
APPLICANT: Franchini, Genoveffa
APPLICANT: Markham, Phillip D.
APPLICANT: Gallo, Robert C.
APPLICANT: Lori, Franco C.
APPLICANT: Popovic, Mikulas
APPLICANT: Garner, Suzanne
TITLE OF INVENTION: MOLECULAR CLONES OF HIV-1 AND USES
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
STREET: Eleventh Floor, 1615 L. Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036-5601

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/022,835
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/599,491
FILING DATE: 17-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Scott, Watson T.
REGISTRATION NUMBER: 26,581
REFERENCE/DOCKET NUMBER: WTS/5683/78507/SRL
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 855 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-022-835-6

Query Match
Best Local Similarity 85.1%; Score 2938.5; DB 1; Length 855;
Matches 552; Conservative 27; Mismatches 45; Indels 25; Gaps 3;

2 EKLMTVVYGVVPMKEATTLFCASADAKAYDTEVHNWATACVPTDPNPOEVLNVT 61
31 EKLMTVVYGVVPMKEATTLFCASADAKAYDTEVHNWATACVPTDPNPOEVLNVT 90
62 HNNMKNMNVEMOEDIIISLMDQSLKPCVKLTPLCVTLNCKDV-----NATNTNDSG 115
91 NNNMKNMNVEMOEDIIISLMDQSLKPCVKLTPLCVTLNCKDV-----NATNTNDSG 150


```

QY 116 TMRGEIKNCSFNITTSIRDEVOKEYALFYKLDVXIDNN-NTSYRLISCDTSVITQACP 174
D 151 MVGGEMKNCSFNITTSIRGVQKEYALFYKLDIAPIDNNSNNRYRLISCVITQACP 210
QY 175 KISFEPIPIHYCAPAGFAILKCNKDTNGKAPCKNVSTXCTGIRPVSTOLLNLSLA 234
D 211 KVSFEPIPIHYCAPAGFAILKCKDKKFKNGKPCNTVSTVCTGIRPVSTOLLNLSLA 270
QY 235 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIR 276
D 271 EEEVIRSANPADNAKTIIVOLNESVEINCTRPNNNRKSIHIGPRAFTTGTGIDIR 330
QY 277 QAHCNISRAKNDTLKOIVIKLREOFENKTIIVFNHSSGDEPEIVMHSFNCGEFFYCNST 336
D 331 QAHCNLSRAKNDTLNKIVIKLREOFENKTIIVFKHSSGDEPEIVMHSFNCGEFFYCNST 390
QY 337 QLFNSTNNNTSGSNTEGNTITLPCRKOIINMVEGKAMVAPPIRGQIRCSNITGL 396
D 391 QLFNSTNNNTSGSNTEGNTITLPCRKOIINMVEGKAMVAPPIRGQIRCSNITGL 450
QY 397 LITRDGGINENGTEIFRPGGDMRDNRSSELYKRVVTEIPLGVAFTCKRVRVOREKRA 456
D 451 LITRDGGEDDKTEIFRPGGDMRDNRSSELYKRVVTEIPLGVAFTCKRVRVOREKRA 510
QY 457 VGIGAVELGFLGAAGSTMGAAAMTLTVQARLLSGIVQOQNNLLRAIEAQOHLQLTVWG 516
D 511 VGIGAVELGFLGAAGSTMGAAAMTLTVQARLLSGIVQOQNNLLRAIEAQOHLQLTVWG 570
QY 517 IKOLQARVLAVERYLGDQQLLGWCGSGKLICTTAVPWNASWSKSLDRINNMTWMEWE 576
D 571 IKOLQARVLAVERYLGDQQLLGWCGSGKLICTTAVPWNASWSKSLDRINNMTWMEWE 630
QY 577 REINNTSEIYTLTLEESONOEKNEQELLEDKMASLNNWFDITKMLMY 625
D 631 REINNTSIYTLTLEESONOEKNEQELLEDKMASLNNWFDITKMLMY 679

RESULT 5
US-08-388-809-6
; Sequence 6, Application US/08388809
; Patent No. 5576000
; GENERAL INFORMATION:
; APPLICANT: REITZ, JR., MARVIN S.; FRANCHINI,
; APPLICANT: GENOVEFFA, MARKHAM, PHILIP D. GALLO, ROBERT
; APPLICANT: C.; LORI, FRANCO C.; POPOVIC, MIKULAS; AND
; APPLICANT: GARTNER, SUZANNE
; TITLE OF INVENTION: MOLECULAR CLONES OF HIV-1
; NUMBER OF INVENTION: AND USES THEREOF
; CORRESPONDENCE ADDRESS:
; ADDRESS: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK, 3.5"
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/388,809
; FILING DATE: 15-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/022,835
; FILING DATE: 25-FEB-1993
; APPLICATION NUMBER: US 07/599,491
; FILING DATE: 17-OCT-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: LESLIE A. SERUNIAN
; REGISTRATION NUMBER: 35,353
; REFERENCE/DOCKET NUMBER: 2026-4092US2

```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 855 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-388-809-6

Query Match 87.7%; Score 2938.5; DB 1; Length 855;
Best Local Similarity 85.1%; Pred. No. 3,4e-245;
Matches 552; Conservative 27; Mismatches 45; Indels 25; Gaps 3;

QY 2 EKLMTVYGVVWPKKETTTLTFCASDAKAYDEVTHANNMATHACVPTDPNQEVLNVTE 61
D 31 EKLMTVYGVVWPKKETTTLTFCASDAKAYDEVTHANNMATHACVPTDPNQEVLNVTE 90
QY 62 HFNWKNMVEQOEEDIIISIMDOSLKPCVKLTPLCVTLNCKDV-----NATNTDSEG 115
D 91 HFNWKNMVEQOEEDIIISIMDOSLKPCVKLTPLCVTLNCKDLNATNGDNTTSSRG 150
QY 116 TMRGEIKNCSFNITTSIRDEVOKEYALFYKLDVXIDNN-NTSYRLISCDTSVITQACP 174
D 151 MVGGEMKNCSFNITTSIRGVQKEYALFYKLDIAPIDNNSNNRYRLISCVITQACP 210
QY 175 KISFEPIPIHYCAPAGFAILKCNKDTNGKAPCKNVSTXCTGIRPVSTOLLNLSLA 234
D 211 KVSFEPIPIHYCAPAGFAILKCKDKKFKNGKPCNTVSTVCTGIRPVSTOLLNLSLA 270
QY 235 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIR 276
D 271 EEEVIRSANPADNAKTIIVOLNESVEINCTRPNNNRKSIHIGPRAFTTGTGIDIR 330
QY 277 QAHCNISRAKNDTLKOIVIKLREOFENKTIIVFNHSSGDEPEIVMHSFNCGEFFYCNST 336
D 331 QAHCNLSRAKNDTLNKIVIKLREOFENKTIIVFKHSSGDEPEIVMHSFNCGEFFYCNST 390
QY 337 QLFNSTNNNTSGSNTEGNTITLPCRKOIINMVEGKAMVAPPIRGQIRCSNITGL 396
D 391 QLFNSTNNNTSGSNTEGNTITLPCRKOIINMVEGKAMVAPPIRGQIRCSNITGL 450
QY 397 LITRDGGINENGTEIFRPGGDMRDNRSSELYKRVVTEIPLGVAFTCKRVRVOREKRA 456
D 451 LITRDGGEDDKTEIFRPGGDMRDNRSSELYKRVVTEIPLGVAFTCKRVRVOREKRA 510
QY 457 VGIGAVELGFLGAAGSTMGAAAMTLTVQARLLSGIVQOQNNLLRAIEAQOHLQLTVWG 516
D 511 VGIGAVELGFLGAAGSTMGAAAMTLTVQARLLSGIVQOQNNLLRAIEAQOHLQLTVWG 570
QY 517 IKOLQARVLAVERYLGDQQLLGWCGSGKLICTTAVPWNASWSKSLDRINNMTWMEWE 576
D 571 IKOLQARVLAVERYLGDQQLLGWCGSGKLICTTAVPWNASWSKSLDRINNMTWMEWE 630
QY 577 REINNTSEIYTLTLEESONOEKNEQELLEDKMASLNNWFDITKMLMY 625
D 631 REINNTSIYTLTLEESONOEKNEQELLEDKMASLNNWFDITKMLMY 679

RESULT 6
US-08-647-714-6
; Sequence 6, Application US/08647714
; Patent No. 5869313
; GENERAL INFORMATION:
; APPLICANT: REITZ, JR., MARVIN S.; FRANCHINI,
; APPLICANT: GENOVEFFA, MARKHAM, PHILIP D. GALLO, ROBERT
; APPLICANT: C.; LORI, FRANCO C.; POPOVIC, MIKULAS; AND
; APPLICANT: GARTNER, SUZANNE
; TITLE OF INVENTION: MOLECULAR CLONES OF HIV-1
; NUMBER OF INVENTION: AND USES THEREOF

```

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: MORGAN & PINNEGAN
;; STREET: 345 PARK AVENUE
;; CITY: NEW YORK
;; STATE: NEW YORK
;; COUNTRY: USA
;; ZIP: 10154
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: FLOPPY DISK, 3.5"
;; COMPUTER: IBM PC COMPATIBLE
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: WORDPERFECT 5.1
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/647,714
;; FILING DATE:
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: US/08/388,809
;; FILING DATE: 15-FEB-1995
;; APPLICATION NUMBER: US 08/022,835
;; FILING DATE: 25-FEB-1993
;; APPLICATION NUMBER: US 07/599,491
;; FILING DATE: 17-OCT-1990
;; ATTORNEY/AGENT INFORMATION:
;; NAME: LESLIE A. SERUNIAN
;; REGISTRATION NUMBER: 35,353
;; REFERENCE/DOCKET NUMBER: 2026-4092US2
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 758-4800
;; TELEFAX: (212) 751-6849
;;
;; INFORMATION FOR SEQ ID NO: 6:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 855 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-647-714-6

Query Match 87.7%; Score 2938.5; DB 2; Length 855;
Best Local Similarity 85.1%; Pred. No. 3.4e-245;
Matches 552; Conservative 27; Mismatches 45; Indels 25; Gaps 3;

QY 2 EKLWTVVYGVVPMWEATTTLCASDAKAYDTEVHNWATHACVPTDNPQEVLENTYE 61
DB 31 EKLWTVVYGVVPMWEATTTLCASDAKAYDTEVHNWATHACVPTDNPQEVLENTYE 90
QY 62 HNNMKNMNVEMQMDIISLDQSLKPCVKLTPLCVTLNCDV-----NATNTTNDSEG 115
DB 91 HNNMKNMNVEMQMDIISLDQSLKPCVKLTPLCVTLNCDV-----NATNTTNDSEG 150
QY 116 TWERGEIKNCSFNITTSIRDEVOKEVALFYKLDVYXIDNN-NTSVRLISCDTSVITQACP 174
DB 151 WGGGBMKKCSFNITTNIRKQVKEVALFYKLDIADIDNNSNRRLISCDTSVITQACP 210
QY 175 KISFEPIPIHYCAPAGFALIKCNKDTFNGKXPCKNVSTYQCTHGIRPVVSTOLLNGSLA 234
DB 211 KVSFEPIPIHYCAPAGFALIKCNKDKKFNKGTNVSTYQCTHGIRPVVSTOLLNGSLA 270
QY 235 EEEVVIIRSDNFTNNAKTIIIVOLKESEVEINCTRPNNNG-----AGDIR 276
DB 271 EEEVVIIRSDNFTNNAKTIIIVOLKESEVEINCTRPNNNG-----AGDIR 330
QY 277 QAHCHISRAKNDTKQOIVIKLREOPENKTIVFNHSSGGDEPILVMSFNGGSEPFYCNST 336
DB 331 QAHCHISRAKNDTKQOIVIKLREOPENKTIVFKHSSGGDEPILVMSFNGGSEPFYCNST 390
QY 337 QLFNSTWNNNTGSSNNTGNTTTLPCRKOIIMMQEVEKAMVAPPIRQIICSSNITGL 396
DB 391 QLFNSTWNNNTGSSNNTGNTTTLPCRKOIIMMQEVEKAMVAPPIRQIICSSNITGL 450
QY 397 LLTRGGGNGENGTLEFRPGGGMDRDMWSELKYKVKYKEPIGVAATKCKRRVVOREKKA 456
DB 451 LLTRGGGNGENGTLEFRPGGGMDRDMWSELKYKVKYKEPIGVAATKCKRRVVOREKKA 510

QY 457 VGIQAVFLGFLGAGSTWGAASMTLTVQARLLSGIVQOQNNLLRAIEAQRMLOLTWVG 516
DB 511 VGIQAVFLGFLGAGSTWGAASMTLTVQARLLSGIVQOQNNLLRAIEAQRMLOLTWVG 570
QY 517 IKOLQARVLAVERYIGDOQLGICWCGSKLTCTAVPNNASVSNKSLDIRNNNTWMEWE 576
DB 571 IKOLQARVLAVERYIGDOQLGICWCGSKLTCTAVPNNASVSNKSLDIRNNNTWMEWE 630
QY 577 REINDVTSEIYTLIESONQOEKNEOPLLELDKASLNNWPDITKIMLY 625
DB 631 REINDVTSEIYTLIESONQOEKNEOPLLELDKASLNNWPDITKIMLY 679

RESULT 7
US-07-956-483-14
; Sequence 14, Application US/07956483
; Patent No. 6261799
; GENERAL INFORMATION:
; APPLICANT: KIENY, Marie-Paule
; TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P. O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22113-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/956,483
; FILING DATE: 31-DEC-1992
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: WO 92/19742
; FILING DATE: 12-NOV-1992
; APPLICATION INFORMATION:
; APPLICATION NUMBER: FR 91 05392
; FILING DATE: 02-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Crane-Feuilly, Sharon E
; REGISTRATION NUMBER: 36,113
; REFERENCE/DOCKET NUMBER: 01753-005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 855 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-07-956-483-14

Query Match 86.6%; Score 2903.5; DB 3; Length 855;
Best Local Similarity 84.4%; Pred. No. 3.6e-242;
Matches 553; Conservative 31; Mismatches 36; Indels 35; Gaps 8;

QY 2 EKLWTVVYGVVPMWEATTTLCASDAKAYDTEVHNWATHACVPTDNPQEVLENTYE 61
DB 31 EKLWTVVYGVVPMWEATTTLCASDAKAYDTEVHNWATHACVPTDNPQEVLENTYE 90
QY 62 HNNMKNMNVEMQMDIISLDQSLKPCVKLTPLCVTLNCDV-----NATNTTNDSEG 115
DB 91 HNNMKNMNVEMQMDIISLDQSLKPCVKLTPLCVTLNCDV-----NATNTTNDSEG 150

```

Qy 116 TMRGEIKNCSFNITTSIRDEVOKEYALFYKLDVXIDNNNTSYRLISCDTSVITQACP 175
Db 151 KMEGEMTNCSFNITTSIRSKQKEYALFYKLDVVPID--NTSYLINCNTSVITQACP 208
Qy 176 ISFEPPIHYCAPAGFALIKNDKTFNGKAPCKNVSTVQCTHGRPVVSTOLLNGSLAE 235
Db 209 VSFEPPIHYCA--RWFALINCNNKFNCTGCTVNSTVQCTHGRPVVSTHLLNGSLAE 267
Qy 236 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNGA-----GDIRQ 277
Db 268 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNTTSHIHGPRAFYATGDIIGDIRQ 327
Qy 278 AHCNISRAKNDTLKQIYIKLEQFENKTIVFNHSSGDPETVHNSFCGGEFFYCNSQ 337
Db 328 AHCNISRAKNDTLKQIYIKLEQFENKTIIIFNRSAGDPETVHNSFCGGEFFYCNSQ 387
Qy 338 LFNSTNNNTSGSNNTBEN--TTLPCRKQIINNMOEYKAMVAPPIGQIRCSNITGL 396
Db 388 LFNSTNNNTSGSNNTBEN--TTLPCRKQIINNMOEYKAMVAPPIGQIRCSNITGL 446
Qy 397 LITRDGGINENG-----TEIFRPGGDMRDNRSELKYKVKYKIEPLGVAFTKCKRRVVO 451
Db 447 LITRDGGINENG-----TEIFRPGGDMRDNRSELKYKVKYKIEPLGVAFTKCKRRVVO 506
Qy 452 REKRAVG--IGAIFLGAAGSTWGAASMTLTVQARLLLSGIVQOQNNLLRAIEAQORML 510
Db 507 REKRAVG--IGAIFLGAAGSTWGAASMTLTVQARLLLSGIVQOQNNLLRAIEAQORML 566
Qy 511 QLTWGIKQOLARVLAVERVYLDQDLGIGCSGKLTCTTAVPNMASHNSKSLDRINNM 570
Db 567 QLTWGIKQOLARVLAVERVYLDQDLGIGCSGKLTCTTAVPNMASHNSKSLDRINNM 626
Qy 571 TWMEREIDNVTSEIYTLIESQNOEKNOEELLEDKMASLWNWPDITKMLWY 625
Db 627 TWMEREIDNVTSEIYTLIESQNOEKNOEELLEDKMASLWNWPDITKMLWY 681

```

RESULT 8
US-08-127-499A-14
Sequence 14, Application US/08127499A
Patent No. 510264

GENERAL INFORMATION:
APPLICANT: VAN ALSTYNE, Diane
TITLE OF INVENTION: ANTIBODIES WHICH BIND MENINGITIS RELATED
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/127.499A
FILING DATE: 28-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 51916/102/INBI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136

INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 861 amino acids

TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
US-08-127-499A-14

Query Match 86.6%; Score 2901.5; DB 1; Length 861;
Best Local Similarity 84.3%; Pred. No. 5.5e-242;
Matches 552; Conservative 31; Mismatches 41; Indels 31; Gaps 7;

```

Qy 2 EKLWVTVYGVVWKEATTTTLFCASDAKAYDTEVHNVMAHACVPTDPNPOEVLEANTE 61
Db 32 EKLWVTVYGVVWKEATTTTLFCASDAKAYDTEVHNVMAHACVPTDPNPOEVLEANTE 91
Qy 62 HNNMKNMWEQWQEDIIISLMDQSLKPCVLTPLCVTLNCKDV--NATNT---TNDSEG- 115
Db 92 NFNMMKNDWVEQWHEDIISLMDQSLKPCVLTPLCVSLKCTLCGNATNTNSNTSSGE 151
Qy 116 -TMRGEIKNCSFNITTSIRDEVOKEYALFYKLDVXIDNNNTSYRLISCDTSVITQACP 174
Db 152 MMEKGEIKNCSFNITTSIRDEVOKEYALFYKLDIIPIDDTTSYTLSCNTSVITQACP 211
Qy 175 KISFEPPIHYCAPAGFALIKNDKTFNGKAPCKNVSTVQCTHGRPVVSTOLLNGSLA 234
Db 212 KISFEPPIHYCAPAGFALIKNDKTFNGKAPCKNVSTVQCTHGRPVVSTOLLNGSLA 271
Qy 235 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNN-----GAGDI 275
Db 272 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNRKSIIRIGRGRAPVTIGKIGNM 331
Qy 276 ROAHNISRAKNDTLKQIYIKLEQF--ENKTIVFNHSSGDPETVHNSFCGGEFFYC 334
Db 332 ROAHNISRAKNDTLKQIYIKLEQF--ENKTIVFNHSSGDPETVHNSFCGGEFFYC 391
Qy 335 STOLFNSSTWNTS--TNGSNNTBS--NTTLPCRKQIINNMOEYKAMVAPPIGQIRCS 390
Db 392 STOLFNSSTWNTS--TNGSNNTBS--NTTLPCRKQIINNMOEYKAMVAPPIGQIRCS 451
Qy 391 SNITGLLITRDGGINENGTEIFRPGGDMRDNRSELKYKVKYKIEPLGVAFTKCKRRV 450
Db 452 SNITGLLITRDGGINENGTEIFRPGGDMRDNRSELKYKVKYKIEPLGVAFTKCKRRV 511
Qy 451 QREKRAVG--IGAIFLGAAGSTWGAASMTLTVQARLLLSGIVQOQNNLLRAIEAQORML 510
Db 512 QREKRAVG--IGAIFLGAAGSTWGAASMTLTVQARLLLSGIVQOQNNLLRAIEAQORML 571
Qy 511 QLTWGIKQOLARVLAVERVYLDQDLGIGCSGKLTCTTAVPNMASHNSKSLDRINNM 570
Db 572 QLTWGIKQOLARVLAVERVYLDQDLGIGCSGKLTCTTAVPNMASHNSKSLDRINNM 631
Qy 571 TWMEREIDNVTSEIYTLIESQNOEKNOEELLEDKMASLWNWPDITKMLWY 625
Db 632 TWMEREIDNVTSEIYTLIESQNOEKNOEELLEDKMASLWNWPDITKMLWY 686

```

RESULT 9
US-08-482-847-14
Sequence 14, Application US/08482847
Patent No. 5556757

GENERAL INFORMATION:
APPLICANT: VAN ALSTYNE, Diane
TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPIC SITES FOR
TITLE OF INVENTION: BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/482,847
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/127,499
 FILING DATE: 28-SEP-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: BENT, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 51916/104/INBI
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202)672-5300
 TELEFAX: (202)672-5399
 TELEX: 904136
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 861 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: unknown
 US-08-482-847-14

Query Match 86.6%; Score 2901.5; DB 1; Length 861;

Best Local Similarity 84.3%; Pred. No. 5.5e-242; Matches 552; Conservative 31; Mismatches 41; Indels 31; Gaps 7;

2 EKLMTVVYGVVPMKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTE 61
 32 EKLMTVVYGVVPMKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTE 91
 62 HNNMKNMVEQMOEDIIISLMDQSLKPCVKLTPLCVTLNCKDV-NATNT---TNDSEG- 115
 92 NNNMKNMVEQMOEDIIISLMDQSLKPCVKLTPLCVTLNCKDV-NATNT---TNDSEG- 151
 116 -TMRGEIKNCSFNITTSIRDEVQKEXALFYKLDVYXIDNNNTSVRLSCDTSVITQACP 174
 152 MMEKGEIKNCSFNITTSIRDEVQKEXALFYKLDVYXIDNNNTSVRLSCDTSVITQACP 211
 175 KISFEPIPIHYCAPAGFALIKCNDKTFNGKPCPKNVSTXQCTHGIRPVVSTOLLNGSLA 234
 212 KVSFEPIPIHYCAPAGFALIKCNDKTFNGKPCPKNVSTXQCTHGIRPVVSTOLLNGSLA 271
 235 EEEVIRSDNFTNNAKTTIIVOLKESVEINCTRPNNN-----GAGDI 275
 272 EEEVIRSDNFTNNAKTTIIVOLKESVEINCTRPNNN-----GAGDI 331
 276 ROAHNINISAKKNATLKOIVILRQF-BNKTIIVNHSAGSPETVMHSFNGGEPFYCN 334
 332 ROAHNINISAKKNATLKOIVILRQF-BNKTIIVNHSAGSPETVMHSFNGGEPFYCN 391
 335 STOLFNSTWNN---NTEGSNTEG-NTITLPCRIOIIMMOBVGKAMAPPIROIRCS 390
 392 STOLFNSTWNN---NTEGSNTEG-NTITLPCRIOIIMMOBVGKAMAPPIROIRCS 451
 391 SNITGLLLTRDGGINENGTEIFRPGGDMRDMRSELVYKVKI EPLGVAFTKCRVY 450
 452 SNITGLLLTRDGGINENGTEIFRPGGDMRDMRSELVYKVKI EPLGVAFTKCRVY 511
 451 OREKRAVIGAFGLFLAAGSTWGAASVTLTVQARLLSGIVQOONMLRLAIBEQOHL 571
 512 OREKRAVIGAFGLFLAAGSTWGAASVTLTVQARLLSGIVQOONMLRLAIBEQOHL 571
 511 QLTWGIKOLQARVLAVERVLDQOLGIGWSSGKLICTTAVPMNASNKSILDRIMNM 570
 572 QLTWGIKOLQARVLAVERVLDQOLGIGWSSGKLICTTAVPMNASNKSILDRIMNM 631
 571 TMEWEIRIDNTYSEIYTLIEESNOQEKNEOELLELDKMASLWMPDITKMLWY 625

Db 632 TMEWEIRIDNTYSEIYTLIEESNOQEKNEOELLELDKMASLWMPDITKMLWY 686

RESULT 10
 US-07-956-483-10
 Sequence 10, Application US/07956483
 Patent No. 6261799
 GENERAL INFORMATION:
 APPLICANT: KIENY, Marie-Paule
 TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED
 NUMBER OF SEQUENCES: 26
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Burns, Doane, Swecker & Mathis
 STREET: P.O. Box 1404
 CITY: Alexandria
 STATE: Virginia
 COUNTRY: United States
 ZIP: 22313-1404
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/956,483
 FILING DATE: 31-DEC-1992
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO 92/19742
 FILING DATE: 12-NOV-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: FR 91 05392
 FILING DATE: 02-MAY-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Crane-Feury, Sharon E
 REGISTRATION NUMBER: 36,113
 REFERENCE/DOCKET NUMBER: 017753-005
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 836-6620
 TELEFAX: (703) 836-2021
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 861 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 US-07-956-483-10

Query Match 86.6%; Score 2901.5; DB 3; Length 861;

Best Local Similarity 84.3%; Pred. No. 5.5e-242; Matches 552; Conservative 31; Mismatches 41; Indels 31; Gaps 7;

2 EKLMTVVYGVVPMKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTE 61
 32 EKLMTVVYGVVPMKEATTTTFCASDAKAYDTEVHNWATHACVPTDNPQEVLENTE 91
 62 HNNMKNMVEQMOEDIIISLMDQSLKPCVKLTPLCVTLNCKDV-NATNT---TNDSEG- 115
 92 NNNMKNMVEQMOEDIIISLMDQSLKPCVKLTPLCVTLNCKDV-NATNT---TNDSEG- 151
 116 -TMRGEIKNCSFNITTSIRDEVQKEXALFYKLDVYXIDNNNTSVRLSCDTSVITQACP 174
 152 MMEKGEIKNCSFNITTSIRDEVQKEXALFYKLDVYXIDNNNTSVRLSCDTSVITQACP 211
 175 KISFEPIPIHYCAPAGFALIKCNDKTFNGKPCPKNVSTXQCTHGIRPVVSTOLLNGSLA 234
 212 KVSFEPIPIHYCAPAGFALIKCNDKTFNGKPCPKNVSTXQCTHGIRPVVSTOLLNGSLA 271
 235 EEEVIRSDNFTNNAKTTIIVOLKESVEINCTRPNNN-----GAGDI 275
 272 EEEVIRSDNFTNNAKTTIIVOLKESVEINCTRPNNN-----GAGDI 331

```

QY 276 ROAHCHISRAKNDTLKQIVIKLREOF-ENKTIIVFNHSSGGDEIIVHSHFNCGEFFYCN 334
D 332 ROAHCHISRAKNDTLKQIVIKLREOF-ENKTIIVFNHSSGGDEIIVHSHFNCGEFFYCN 391
QY 335 STOLFNSWTNN--NTEGSNTEG-NTITLPCRIKQIINMOVGVKAMVAPPIRGQIRCS 390
D 392 STOLFNSWTNN--NTEGSNTEG-NTITLPCRIKQIINMOVGVKAMVAPPIRGQIRCS 451
QY 391 SNITGLLTBDGNGENGTIFRPGGDMRNMSELYKKVVKIEPLGVAFTCKRRV 450
D 452 SNITGLLTBDGNGENGTIFRPGGDMRNMSELYKKVVKIEPLGVAFTCKRRV 511
QY 451 QREKRAVIGAVFLGFGAAGSTMGASMTLTVQARLLSGIYOQONNLRAIEAOQRL 510
D 512 QREKRAVIGAVFLGFGAAGSTMGASMTLTVQARLLSGIYOQONNLRAIEAOQRL 571
QY 511 QLTWGIKQLOARLAVERYLGDQOLGIGWCSGKLICTTAVPWNASWSKSLDRINNM 570
D 572 QLTWGIKQLOARLAVERYLGDQOLGIGWCSGKLICTTAVPWNASWSKSLDRINNM 631
QY 571 TMMEREIDNYTSEIYTLIESONOEKNEOELLEDKVASLNNFDTIKMLY 625
D 632 TMMEREIDNYTSEIYTLIESONOEKNEOELLEDKVASLNNFDTIKMLY 686

```

RESULT 11

```

US-08-472-240A-1
Sequence 1, Application US/08472240A
Patent No. 6284248
GENERAL INFORMATION:
APPLICANT: KIENY, Marie-Paule
TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,240A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/956,483
FILING DATE: 31-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-055
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 661 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Proteiin
LOCATION: 1..854
US-08-472-240A-1

```

Query Match 86.6%; Score 2901.5; DB 3; Length 861;
 Best Local Similarity 84.3%; Pred. No. 5,5e-242;
 Matches 552; Conservative 31; Mismatches 41; Indels 31; Gaps 7;

```

QY 2 EKLMTVYGVVWKEATTTLPQASDAKAYDTEVHNVMAHACVPPDPNQEVLNATE 61
D 32 EKLMTVYGVVWKEATTTLPQASDAKAYDTEVHNVMAHACVPPDPNQEVLNATE 91
QY 62 HFMWKNWVEQOEEDISLMDOSLKPVCVLTPLCTLNCKDV-NATNT-----TNDEG- 115
D 92 HFMWKNWVEQOEEDISLMDOSLKPVCVLTPLCTLNCKDV-NATNT-----TNDEG- 151
QY 116 -TMRGEIKNCSFNITTSIRDEVEQVYALFYKLDVYXIDNNNTSYRLISCDTSVITQACP 174
D 152 MMKEKEIKNCSFNITTSIRDEVEQVYALFYKLDVYXIDNNNTSYRLISCDTSVITQACP 211
QY 175 KISPEPIPIHYCAPAGPAILIKNDKTPNGCKPCKNSTXCTGTGIRPVSTOLLNGSLA 234
D 212 KISPEPIPIHYCAPAGPAILIKNDKTPNGCKPCKNSTXCTGTGIRPVSTOLLNGSLA 271
QY 235 EEEVIRSDNFTNNAKTIIVQLKESVEINCTRPNN-----GAGDI 275
D 272 EEEVIRSDNFTNNAKTIIVQLKESVEINCTRPNN-----GAGDI 331
QY 276 ROAHCHISRAKNDTLKQIVIKLREOF-ENKTIIVFNHSSGGDEIIVHSHFNCGEFFYCN 334
D 332 ROAHCHISRAKNDTLKQIVIKLREOF-ENKTIIVFNHSSGGDEIIVHSHFNCGEFFYCN 391
QY 335 STOLFNSWTNN--NTEGSNTEG-NTITLPCRIKQIINMOVGVKAMVAPPIRGQIRCS 390
D 392 STOLFNSWTNN--NTEGSNTEG-NTITLPCRIKQIINMOVGVKAMVAPPIRGQIRCS 451
QY 391 SNITGLLTBDGNGENGTIFRPGGDMRNMSELYKKVVKIEPLGVAFTCKRRV 450
D 452 SNITGLLTBDGNGENGTIFRPGGDMRNMSELYKKVVKIEPLGVAFTCKRRV 511
QY 451 QREKRAVIGAVFLGFGAAGSTMGASMTLTVQARLLSGIYOQONNLRAIEAOQRL 510
D 512 QREKRAVIGAVFLGFGAAGSTMGASMTLTVQARLLSGIYOQONNLRAIEAOQRL 571
QY 511 QLTWGIKQLOARLAVERYLGDQOLGIGWCSGKLICTTAVPWNASWSKSLDRINNM 570
D 572 QLTWGIKQLOARLAVERYLGDQOLGIGWCSGKLICTTAVPWNASWSKSLDRINNM 631
QY 571 TMMEREIDNYTSEIYTLIESONOEKNEOELLEDKVASLNNFDTIKMLY 625
D 632 TMMEREIDNYTSEIYTLIESONOEKNEOELLEDKVASLNNFDTIKMLY 686

```

RESULT 12

```

US-08-472-240A-7
Sequence 7, Application US/08472240A
Patent No. 6284248
GENERAL INFORMATION:
APPLICANT: KIENY, Marie-Paule
TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,240A
FILING DATE: 07-JUN-1995

```

```

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/956,483
FILING DATE: 31-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-055
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 861 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Protein
LOCATION: 1..831
DB-472-240A-7

```

```

Query Match      86.6%; Score 2901.5; DB 3; Length 861;
Best Local Similarity 84.3%; Pred. No. 5.5e-242;
Matches 552; Conservative 31; Mismatches 41; Indels 31; Gaps 7;

```

```

QY 2 EKLAWTVVYGVVPMKEATTTTFCASDAKAYDTEVHNWATACVPTDNPPOEVLENTE 61
DB 32 EKLAWTVVYGVVPMKEATTTTFCASDAKAYDTEVHNWATACVPTDNPPOEVLENTE 91
QY 62 HNNMKNMNVQWQEDIIISLWDSLKPCVKLTPLCVTLNCKDV-NATNT---TNDSEG- 115
DB 92 NFNMMKNDVQWQEDIIISLWDSLKPCVKLTPLCVSLKCTDLGNATNTNSNTSSSGE 151
QY 116 -TMEGEIKNSFNITTSIRDEVQKEYALFYKLDVXVXIDNNNTSYRLISCSTSVITQACP 174
DB 152 MMEKEGEEKNSFNISTIRGVOKEYAFYKLDIIPIDNDTTSYRLISCSTSVITQACP 211
QY 175 KISFEPIPIHYCAPGAFALCKNDKTFNGKPKCKNVSTQCTHGIRPVVSTOLLNGSLA 234
DB 212 KISFEPIPIHYCAPGAFALCKNDKTFNGKPKCKNVSTQCTHGIRPVVSTOLLNGSLA 271
QY 235 EEEVVISRDNFTNNAKTIIIVOLKESVEINCTRPNNN-----GAGDI 275
DB 272 EEEVVISRDNFTNNAKTIIIVOLKESVEINCTRPNNNTRKSIIRQPGRAFVTIGKIGM 331
QY 276 ROAHGNISRAKKNATLKQIVIKLRQF-ENKTIIVNHSNGDPEIIVMHSFNGSGEPYCN 334
DB 332 ROAHGNISRAKKNATLKQIVIKLRQF-ENKTIIVNHSNGDPEIIVMHSFNGSGEPYCN 391
QY 335 STOLFNSNTMN---NTEGSSNTEG-NTITLPCRIOIIMQWQKAMVAPPIRQIICS 390
DB 392 STOLFNSNTMN---NTEGSSNTEGSDTITLPCRIOIIMQWQKAMVAPPIRQIICS 451
QY 391 SNITGLLITRDOGIENGTEIFRPGCGDMRDNRSELVYKVKYLEPLGVAATYCKRRVY 450
DB 452 SNITGLLITRDOGINNNGSEIFRPGCGDMRDNRSELVYKVKYLEPLGVAATYCKRRVY 511
QY 451 QREKRAVIGAVLFGAAGSTWGAASWTLTVQARLLLSGIVQOQNNLLRAIEAQOQRL 510
DB 512 QREKRAVIGAVLFGAAGSTWGAASWTLTVQARLLLSGIVQOQNNLLRAIEAQOQRL 571
QY 511 QLTWGIQOLQARVLAVERYLGDQQLLGIGWSSGLICTTAVPMAASNSKSLDIAIWM 570
DB 572 QLTWGIQOLQARVLAVERYLGDQQLLGIGWSSGLICTTAVPMAASNSKSLDIAIWM 631
QY 571 TMEWEERIDNVTSEIYTLIESQNOQEKNEBELLELDKMAIWMFDTTKLWY 625
DB 632 TMEWEERIDNVTSEIYTLIESQNOQEKNEBELLELDKMAIWMFDTTKLWY 686

```

RESULT 13

```

US-08-817-441-103
Sequence 103, Application US/08817441
Patent No. 6399294
GENERAL INFORMATION:
APPLICANT: CHARNEAU, PIERRE
APPLICANT: CLAVEL, FRANCOISE
APPLICANT: BORMAN, ANDREW
APPLICANT: OUILLENT, CAROLINE
APPLICANT: GUETARD, DENISE
APPLICANT: MONTAGNIER, LUC
APPLICANT: DONJON DE SAINT-MARTIN, JOCELINE
TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF HIV-1 TYPE (OR
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner, L.L.P.
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/817,441
FILING DATE: 11-JUL-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR 95/01391
FILING DATE: 20-OCT-1995
APPLICATION NUMBER: FR 9412554
FILING DATE: 20-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9502526
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03260, 6005-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 103:
SEQUENCE CHARACTERISTICS:
LENGTH: 861 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-817-441-103
Query Match      86.6%; Score 2901.5; DB 4; Length 861;
Best Local Similarity 84.3%; Pred. No. 5.5e-242;
Matches 552; Conservative 31; Mismatches 41; Indels 31; Gaps 7;

```

```

QY 2 EKLAWTVVYGVVPMKEATTTTFCASDAKAYDTEVHNWATACVPTDNPPOEVLENTE 61
DB 32 EKLAWTVVYGVVPMKEATTTTFCASDAKAYDTEVHNWATACVPTDNPPOEVLENTE 91
QY 62 HNNMKNMNVQWQEDIIISLWDSLKPCVKLTPLCVTLNCKDV-NATNT---TNDSEG- 115
DB 92 NFNMMKNDVQWQEDIIISLWDSLKPCVKLTPLCVSLKCTDLGNATNTNSNTSSSGE 151
QY 116 -TMEGEIKNSFNITTSIRDEVQKEYALFYKLDVXVXIDNNNTSYRLISCSTSVITQACP 174
DB 152 MMEKEGEEKNSFNISTIRGVOKEYAFYKLDIIPIDNDTTSYRLISCSTSVITQACP 211
QY 175 KISFEPIPIHYCAPGAFALCKNDKTFNGKPKCKNVSTQCTHGIRPVVSTOLLNGSLA 234

```

```

Db 212 KVSFEPIPIHYCAPAGFAILKCNNTFNGTGPCTVSTVQCTHGIRPVSTOLLNGSLA 271
Qy 235 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNN-----GAGDI 275
Db 272 EEEVIRSANFTDAKTIIVOLNOSVEINCTRPNNNRKSIIRIORGPRAFVITGKIGNM 331
Qy 276 ROACHNISRAKNDTLKOIIVIKLEOF-ENKTIIVFNHSSGDDPEIVHMSFNGGEPFYCN 334
Db 332 ROACHNISRAKNDTLKOIIVIKLEOF-ENKTIIVFNHSSGDDPEIVHMSFNGGEPFYCN 391
Qy 335 STOLFNSTWNN--NTGSSNTEG-NTITLPCRKOIINMOEYKAMVAPPISGQIRCS 390
Db 392 STOLFNSTWNN--NTGSSNTEG-NTITLPCRKOIINMOEYKAMVAPPISGQIRCS 451
Qy 391 SNITGLLLTRDGINENGTETFRPGGDMRDNMRSELYKRVKIEPLGVAFTAKRRV 450
Db 452 SNITGLLLTRDGINENGTETFRPGGDMRDNMRSELYKRVKIEPLGVAFTAKRRV 511
Qy 451 OREKRAVGIGAVFLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOONNLLRAIEAQRM 510
Db 512 OREKRAVGIGAVFLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOONNLLRAIEAQRM 571
Qy 511 QLTWGIKQOLARILAVERYLGDQQLGIWCGSKLICTTAVPNNASWSNKSLEDRIMNM 570
Db 572 QLTWGIKQOLARILAVERYLGDQQLGIWCGSKLICTTAVPNNASWSNKSLEDRIMNM 631
Qy 571 TWMEBERIDNVTSEIYTLIESONQOEKNEOELLEDKNASLNMWFDITKLMY 625
Db 632 TWMEBERIDNVTSEIYTLIESONQOEKNEOELLEDKNASLNMWFDITKLMY 686

```

RESULT 14

```

US-07-956-483-16
; Sequence 16, Application US/07956483
; Patent No. 6261799
; GENERAL INFORMATION:
; APPLICANT: KIENY, Marie-Paule
; TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED
; TITLE OF INVENTION: 9P160 VARIANT
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P. O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/956,483
; FILING DATE: 31-DEC-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 92/19742
; FILING DATE: 12-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91 05392
; FILING DATE: 02-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Crane-Feury, Sharon E
; REGISTRATION NUMBER: 36,113
; REFERENCE/DOCKET NUMBER: 017753-005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 861 amino acids

```

```

; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-07-956-483-16

```

```

Query Match 86.3%; Score 2893.5; DB 3; Length 861;
Best Local Similarity 84.1%; Pred. No. 2,7e-241;
Matches 551; Conservative 31; Mismatches 42; Indels 31; Gaps 7;

```

```

Qy 2 EKLWVTVYGVYVWKEKATITLFCASDAKAYDTEHANNMATHACVPTDPNQEVLNVT 61
Db 3 EKLWVTVYGVYVWKEKATITLFCASDAKAYDTEHANNMATHACVPTDPNQEVLNVT 91
Qy 62 HNNWKNWYEQMEDIIISLDOSLKPVCVLTPLCVTLNCKDV-NATNT---TNDSEG- 115
Db 92 HNNWKNWYEQMEDIIISLDOSLKPVCVLTPLCVTLNCKDV-NATNT---TNDSEG- 151
Qy 116 -TWGEIKKCSFNITTSIRDEVOKEYALFYKLDVXIDNNNTSYRLISGDSYITGACP 174
Db 152 MMMEKGEIKKCSFNITTSIRDEVOKEYALFYKLDVXIDNNNTSYRLISGDSYITGACP 211
Qy 175 KISFEPIPIHYCAPAGFAILKCNNDKTNGKPKCNVSTVQCTHGIRPVSTOLLNGSLA 234
Db 212 KVSFEPIPIHYCAPAGFAILKCNNDKTNGKPKCNVSTVQCTHGIRPVSTOLLNGSLA 271
Qy 235 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNN-----GAGDI 275
Db 272 EEEVIRSANFTDAKTIIVOLNOSVEINCTRPNNNRKSIIRIORGPRAFVITGKIGNM 331
Qy 276 ROACHNISRAKNDTLKOIIVIKLEOF-ENKTIIVFNHSSGDDPEIVHMSFNGGEPFYCN 334
Db 332 ROACHNISRAKNDTLKOIIVIKLEOF-ENKTIIVFNHSSGDDPEIVHMSFNGGEPFYCN 391
Qy 335 STOLFNSTWNN--NTGSSNTEG-NTITLPCRKOIINMOEYKAMVAPPISGQIRCS 390
Db 392 STOLFNSTWNN--NTGSSNTEG-NTITLPCRKOIINMOEYKAMVAPPISGQIRCS 451
Qy 391 SNITGLLLTRDGINENGTETFRPGGDMRDNMRSELYKRVKIEPLGVAFTAKRRV 450
Db 452 SNITGLLLTRDGINENGTETFRPGGDMRDNMRSELYKRVKIEPLGVAFTAKRRV 511
Qy 451 OREKRAVGIGAVFLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOONNLLRAIEAQRM 510
Db 512 OREKRAVGIGAVFLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOONNLLRAIEAQRM 571
Qy 511 QLTWGIKQOLARILAVERYLGDQQLGIWCGSKLICTTAVPNNASWSNKSLEDRIMNM 570
Db 572 QLTWGIKQOLARILAVERYLGDQQLGIWCGSKLICTTAVPNNASWSNKSLEDRIMNM 631
Qy 571 TWMEBERIDNVTSEIYTLIESONQOEKNEOELLEDKNASLNMWFDITKLMY 625
Db 632 TWMEBERIDNVTSEIYTLIESONQOEKNEOELLEDKNASLNMWFDITKLMY 686

```

RESULT 15

```

US-09-124-900-9
; Sequence 9, Application US/09124900
; Patent No. 6268484
; GENERAL INFORMATION:
; APPLICANT: KATINGER, Hermann
; APPLICANT: BUCHACHER, Andrea
; APPLICANT: ERNST, Wolfgang
; APPLICANT: BALLAUN, Claudia
; APPLICANT: PUTSCHER, Martin
; APPLICANT: TROLOA, Alexandra
; APPLICANT: PREDL, Renate
; APPLICANT: SCHWATZ, Christine
; APPLICANT: KLIMA, Annelies
; APPLICANT: STEINDL, Franz
; APPLICANT: MUSTER, Thomas
; TITLE OF INVENTION: HIV-Vaccines
; FILE REFERENCE: 1939-112P

```


THIS PAGE BLANK (08.70)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 12:32:44 ; Search time 26.0558 Seconds
(without alignments)
4661.192 Million cell updates/sec

Title: US-10-032-162-17

Perfect score: 3351
Sequence: 1 VEKLMVTYYGVVWKRAAT.....EIDKMSAMNPFITKMLMY 625

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

1 number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3345	99.8	625	US-10-032-162-17	Sequence 17, Appl
2	3295	98.3	643	US-10-032-162-13	Sequence 13, Appl
3	2928	87.4	842	US-10-190-435-2	Sequence 2, Appl
4	2928	87.4	842	US-10-241-009-2	Sequence 2, Appl
5	2928	87.4	842	US-10-190-434B-2	Sequence 2, Appl
6	2928	87.4	842	US-10-190-305A-2	Sequence 2, Appl
7	2928	87.4	847	US-09-476-242-2	Sequence 2, Appl
8	2923	87.4	847	US-09-476-242-2	Sequence 15, Appl
9	2901.5	86.6	861	US-10-026-741-103	Sequence 103, App
10	2878	85.9	856	US-09-476-242-1	Sequence 103, App
11	2867	85.6	856	US-10-196-515-11	Sequence 11, Appl
12	2822.5	84.2	726	US-10-196-515-3	Sequence 3, Appl
13	2815.5	84.0	759	US-10-196-515-12	Sequence 12, Appl
14	2793.5	83.4	868	US-09-938-406-1	Sequence 1, Appl
15	2758	82.3	619	US-09-891-609-4	Sequence 4, Appl

16	2758	82.3	646	US-09-891-609-2	Sequence 2, Appl
17	2611.5	77.9	860	US-10-190-435-6	Sequence 6, Appl
18	2611.5	77.9	860	US-10-241-009-6	Sequence 6, Appl
19	2611.5	77.9	860	US-10-190-434B-6	Sequence 6, Appl
20	2611.5	77.9	860	US-10-190-305A-6	Sequence 6, Appl
21	2572	76.8	867	US-10-190-435-3	Sequence 3, Appl
22	2572	76.8	867	US-10-190-435-126	Sequence 126, App
23	2572	76.8	867	US-10-241-009-3	Sequence 3, Appl
24	2572	76.8	867	US-10-190-434B-3	Sequence 3, Appl
25	2572	76.8	867	US-10-190-305A-3	Sequence 3, Appl
26	2568	76.6	855	US-10-369-294-11	Sequence 11, Appl
27	2568	76.6	855	US-10-369-294-12	Sequence 12, Appl
28	2565.5	76.6	803	US-10-190-435-135	Sequence 135, App
29	2564.5	76.5	803	US-10-190-435-134	Sequence 4, Appl
30	2563	76.5	869	US-10-190-435-4	Sequence 4, Appl
31	2563	76.5	869	US-10-241-009-4	Sequence 4, Appl
32	2563	76.5	869	US-10-190-434B-4	Sequence 4, Appl
33	2563	76.5	869	US-10-190-305A-4	Sequence 4, Appl
34	2562	76.5	853	US-10-369-294-13	Sequence 13, Appl
35	2562	76.5	861	US-10-190-435-139	Sequence 139, App
36	2560	76.4	858	US-10-190-435-150	Sequence 150, App
37	2559.5	76.4	870	US-10-190-435-127	Sequence 127, App
38	2553	76.2	870	US-10-190-435-147	Sequence 147, App
39	2552	76.2	862	US-10-190-435-141	Sequence 141, App
40	2551.5	76.1	865	US-10-190-435-140	Sequence 140, App
41	2550.5	76.1	845	US-10-190-435-129	Sequence 129, App
42	2550.5	76.1	845	US-10-190-435-130	Sequence 130, App
43	2546.5	76.0	857	US-10-190-435-138	Sequence 138, App
44	2542.5	75.9	855	US-10-190-435-144	Sequence 144, App
45	2537.5	75.7	859	US-10-190-435-137	Sequence 137, App

ALIGNMENTS

RESULT 1
US-10-032-162-17
; Sequence 17, Application US/10032162
; Publication No. US20030052839A1
GENERAL INFORMATION:
; APPLICANT: BINLEY, JAMES M
; APPLICANT: SCHUELE, NORBERT
; APPLICANT: OLSON, WILLIAM C
; APPLICANT: PAUL, MADON J
; APPLICANT: JOHN, MOORE P
; TITLE OF INVENTION: STABILIZED VIRAL ENVELOPE PROTEINS AND USES THEREOF
; FILE REFERENCE: 2048/59331az
; CURRENT APPLICATION NUMBER: US/10/032,162
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/602,864
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 625
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (151)..(151)
; OTHER INFORMATION: X-UNKNOWN AMINO ACID
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (205)..(205)
; OTHER INFORMATION: X-UNKNOWN AMINO ACID
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (213)..(213)
; OTHER INFORMATION: X-UNKNOWN AMINO ACID
US-10-032-162-17
Query Match 99.8%; Score 3345; DB 15; Length 625;
Best Local Similarity 100.0%; Pred. No. 9.3e-311;

Matches	61%; Conservative	0; Mismatches	8; Indels	18; Gaps	1
Qy	1	VEKLVMTVYGVGVPMWEKATTTLLFCASDAKAYDTEVHNWMAHACYPDPNPOEVLNENT			60
Db	1	VEKLVMTVYGVGVPMWEKATTTLLFCASDAKAYDTEVHNWMAHACYPDPNPOEVLNENT			60
Qy	61	EHFMNMKNWMEQOEEDIIISLMDSLKPCVLTPLCVTLNCKDNVAATNTDSECTMERG			120
Db	61	EHFMNMKNWMEQOEEDIIISLMDSLKPCVLTPLCVTLNCKDNVAATNTDSECTMERG			120
Qy	121	EIKKCSFNITTSIRDEVOKEVALFYKLDVVAIDNNNTSYRLISCDTSVITQACPRISEEP			180
Db	121	EIKKCSFNITTSIRDEVOKEVALFYKLDVVIDNNNTSYRLISCDTSVITQACPRISEEP			180
Qy	181	IPHYCAPAGFALLKCKDKTEFGKXPCKNVSTXOCTHGIRPVVSTOLLNGLAEVEVI			240
Db	181	IPHYCAPAGFALLKCKDKTEFGKXPCKNVSTOCTHGIRPVVSTOLLNGLAEVEVI			240
Qy	241	RSNDFPTNAKTIIVQLKESVEINCTRPNNNG-----AGDIRAHQNI			282
Db	241	RSNDFPTNAKTIIVQLKESVEINCTRPNNNTRKSIHIGPRAFYTTGELIGDIRAHQNI			300
Qy	283	SRAKNDTLKOIVIKLREOFENKTIIVFNHSSGSDPEIVMHSFNCGEFFYCNSTOLFNST			342
Db	301	SRAKNDTLKOIVIKLRQFENKTIIVFNHSSGSDPEIVMHSFNCGEFFYCNSTOLFNST			360
Qy	343	WNNTEGSNNTEGNTTLPCKRIKQIINMWQEVGKAMVAPPIRGQIRGSSNTIGLLITRDG			402
Db	361	WNNTEGSNNTEGNTTLPCKRIKQIINMWQEVGKAMVAPPIRGQIRGSSNTIGLLITRDG			420
Qy	403	GINNGEIEFPPGGDMDMNRSELYKKVKKIEPLGVAPTCKRRVOREKAVGIGAV			462
Db	421	GINNGEIEFPPGGDMDMNRSELYKKVKKIEPLGVAPTCKRRVOREKAVGIGAV			480
Qy	463	FLGFLGAAGSTMGAASTLTVOARLLLSGIVQOQNNLRAIEAQRMQLTWYGIKOLOA			522
Db	481	FLGFLGAAGSTMGAASTLTVOARLLLSGIVQOQNNLRAIEAQRMQLTWYGIKOLOA			540
Qy	523	RLAVERYLGDQQLLGIWCSGKLICTTAVPMNASWSKSLDRINNNMTWMEWEIDNY			582
Db	541	RLAVERYLGDQQLLGIWCSGKLICTTAVPMNASWSKSLDRINNNMTWMEWEIDNY			600
Qy	583	TSEIYTLIEESQNOQEKNEOELLEIDKVASLWNWDIRKMLNY			625
Db	601	TSEIYTLIEESQNOQEKNEOELLEIDKVASLWNWDIRKMLNY			643
RESULT 3					
US-10-190-435-2					
Sequence 2, Application US/10190435					
Publication No. US20030143248A1					
GENERAL INFORMATION:					
APPLICANT: ZUR MEGEDF, Jan					
APPLICANT: BARNETT, Susan W.					
APPLICANT: LIAN, Ying					
APPLICANT: ENGELBRECHT, Susan					
APPLICANT: VAN RENSBURG, Estrelita J.					
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C					
FILE REFERENCE: PPI8133.003 / 2302-18133					
CURRENT APPLICATION NUMBER: US/10/190,435					
CURRENT FILING DATE: 2002-12-30					
NUMBER OF SEQ ID NOS: 319					
SOFTWARE: PatentIn Ver. 2.0					
SEQ ID NO 2					
LENGTH: 842					
TYPE: PRT					
ORGANISM: Artificial Sequence					
FEATURE:					
OTHER INFORMATION: Description of Artificial Sequence: Sp162					
US-10-190-435-2					
Query Match					
87.4%; Score 2928; DB 12; Length 842;					

Best Local Similarity 86.0%; Pred. No. 1.1e-270;
Matches 555; Conservative 28; Mismatches 40; Indels 22; Gaps 5;

```
QY 1 VEKLVNTVYGVVPMWEKATTTLPFCASDAKAYDTEVHWATACVPTDNPDEVLIENT 60
DB 25 VEKLVNTVYGVVPMWEKATTTLPFCASDAKAYDTEVHWATACVPTDNPDEVLIENT 84
QY 61 EHFNNMKNMNVBOMQEDIIISLWDQSLKPCVKLTPLCVTLNCKDV-NATNTTDSSEGTMR 119
DB 85 EHFNNMKNMNVBOMQEDIIISLWDQSLKPCVKLTPLCVTLNCKDV-NATNTTDSSEGTMR 144
QY 120 GEIKNCSFNITTSIDDEVQKAYALFYKLDVYXIDNNNTSYRLISCDTSYITQACPKISPE 179
DB 145 GEIKNCSFNITTSIDDEVQKAYALFYKLDVYXIDNNNTSYRLISCDTSYITQACPKISPE 204
QY 160 PPIHYCAPAGFALLKCNKDTFNGKPCKNVSTXCTGIRPVSTOLLNLSLAEEV 239
DB 205 PPIHYCAPAGFALLKCNKDTFNGKPCKNVSTXCTGIRPVSTOLLNLSLAEEV 264
QY 240 IRSDNFTNNAKTIIVQLKESVEINCTRPNNNG-----AGDIRQAHGN 281
DB 265 IRSENFNDNAKTIIVQLKESVEINCTRPNNNTRKSITIGPGRAFATGDIIGDIRQAHGN 324
QY 282 ISRAKMDTLKQIVIKLBQFENKTIIVFNHSSGDEIYVMSFNGCGEFFYCNSTOLFNS 341
DB 345 ISGEKNNNTLKQIVTKLQAFGNKTIIVFKOSSGDEIYVMSFNGCGEFFYCNSTOLFNS 384
QY 342 TWNNNTBESNTEGNTITLPCRIKQIIMMOWGVGAMVAPPIRGDIRCSSNTTGLLTLD 401
DB 385 TW-NNTIGPNNNG--TTLPCRIKQIIMMOWGVGAMVAPPIRGDIRCSSNTTGLLTLD 442
QY 402 GGIN-ENGTEIFRPGCGDMRDNRSELKYKVKVIEPLGVAPTKCRVQREKAVGIG 460
DB 443 GGEKESNTTEIFRPGCGDMRDNRSELKYKVKVIEPLGVAPTKCRVQREKAVGIG 502
QY 461 AVFLGFLGAGSTWGAASNTLTVOARLLISGIVQOQNNLLRAIEAQORMLQLTWNGIKOL 520
DB 503 AMFLGFLGAGSTWGAASNTLTVOARLLISGIVQOQNNLLRAIEAQORMLQLTWNGIKOL 562
QY 521 QARVLAVERYLQDQDLGIWGSGLICTTAAPNMAASNSKSLDRIMNMNTMWEREID 580
DB 563 QARVLAVERYLQDQDLGIWGSGLICTTAAPNMAASNSKSLDRIMNMNTMWEREID 622
QY 581 NYTSEIYTLIESONQOEKNEBELLDKMASLMMNPDTTKWLMY 625
DB 623 NYTNLIYTLIESONQOEKNEBELLDKMASLMMNPDTTKWLMY 667
```

RESULT 4
US-10-241-009-2
Sequence 2, Application US/10241009
Publication No. US20030170614A1
GENERAL INFORMATION:
APPLICANT: ZUR MEGEDE, Jan
APPLICANT: BARNETT, Susan
APPLICANT: LIAN, Ying
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B
FILE REFERENCE: 2300-1621.21
CURRENT APPLICATION NUMBER: US/10/241,009
CURRENT FILING DATE: 2002-12-13
NUMBER OF SEQ ID NOS: 68
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 842
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: SFI62
US-10-241-009-2

Query Match 87.4%; Score 2928; DB 12; Length 842;
Best Local Similarity 86.0%; Pred. No. 1.1e-270;
Matches 555; Conservative 28; Mismatches 40; Indels 22; Gaps 5;

Matches 555; Conservative 28; Mismatches 40; Indels 22; Gaps 5;

```
QY 1 VEKLVNTVYGVVPMWEKATTTLPFCASDAKAYDTEVHWATACVPTDNPDEVLIENT 60
DB 25 VEKLVNTVYGVVPMWEKATTTLPFCASDAKAYDTEVHWATACVPTDNPDEVLIENT 84
QY 61 EHFNNMKNMNVBOMQEDIIISLWDQSLKPCVKLTPLCVTLNCKDV-NATNTTDSSEGTMR 119
DB 85 EHFNNMKNMNVBOMQEDIIISLWDQSLKPCVKLTPLCVTLNCKDV-NATNTTDSSEGTMR 144
QY 120 GEIKNCSFNITTSIDDEVQKAYALFYKLDVYXIDNNNTSYRLISCDTSYITQACPKISPE 179
DB 145 GEIKNCSFNITTSIDDEVQKAYALFYKLDVYXIDNNNTSYRLISCDTSYITQACPKISPE 204
QY 160 PPIHYCAPAGFALLKCNKDTFNGKPCKNVSTXCTGIRPVSTOLLNLSLAEEV 239
DB 205 PPIHYCAPAGFALLKCNKDTFNGKPCKNVSTXCTGIRPVSTOLLNLSLAEEV 264
QY 240 IRSDNFTNNAKTIIVQLKESVEINCTRPNNNG-----AGDIRQAHGN 281
DB 265 IRSENFNDNAKTIIVQLKESVEINCTRPNNNTRKSITIGPGRAFATGDIIGDIRQAHGN 324
QY 282 ISRAKMDTLKQIVIKLBQFENKTIIVFNHSSGDEIYVMSFNGCGEFFYCNSTOLFNS 341
DB 345 ISGEKNNNTLKQIVTKLQAFGNKTIIVFKOSSGDEIYVMSFNGCGEFFYCNSTOLFNS 384
QY 342 TWNNNTBESNTEGNTITLPCRIKQIIMMOWGVGAMVAPPIRGDIRCSSNTTGLLTLD 401
DB 385 TW-NNTIGPNNNG--TTLPCRIKQIIMMOWGVGAMVAPPIRGDIRCSSNTTGLLTLD 442
QY 402 GGIN-ENGTEIFRPGCGDMRDNRSELKYKVKVIEPLGVAPTKCRVQREKAVGIG 460
DB 443 GGEKESNTTEIFRPGCGDMRDNRSELKYKVKVIEPLGVAPTKCRVQREKAVGIG 502
QY 461 AVFLGFLGAGSTWGAASNTLTVOARLLISGIVQOQNNLLRAIEAQORMLQLTWNGIKOL 520
DB 503 AMFLGFLGAGSTWGAASNTLTVOARLLISGIVQOQNNLLRAIEAQORMLQLTWNGIKOL 562
QY 521 QARVLAVERYLQDQDLGIWGSGLICTTAAPNMAASNSKSLDRIMNMNTMWEREID 580
DB 563 QARVLAVERYLQDQDLGIWGSGLICTTAAPNMAASNSKSLDRIMNMNTMWEREID 622
QY 581 NYTSEIYTLIESONQOEKNEBELLDKMASLMMNPDTTKWLMY 625
DB 623 NYTNLIYTLIESONQOEKNEBELLDKMASLMMNPDTTKWLMY 667
```

RESULT 5
US-10-190-434B-2
Sequence 2, Application US/10190434B
Publication No. US20030194800A1
GENERAL INFORMATION:
APPLICANT: ZUR MEGEDE, Jan
APPLICANT: BARNETT, Susan
APPLICANT: LIAN, Ying
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B
FILE REFERENCE: 2300-1621.20
CURRENT APPLICATION NUMBER: US/10/190,434B
CURRENT FILING DATE: 2002-07-05
NUMBER OF SEQ ID NOS: 68
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 842
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: SFI62
US-10-190-434B-2

Query Match 87.4%; Score 2928; DB 12; Length 842;
Best Local Similarity 86.0%; Pred. No. 1.1e-270;
Matches 555; Conservative 28; Mismatches 40; Indels 22; Gaps 5;

```

Qy 1 VEKLMVTYVYGVPMWKEATITLFCASDAKAYDTEVHNWMAATHACVPTDPNPOEIVLENT 60
Db 25 VEKLMVTYVYGVPMWKEATITLFCASDAKAYDTEVHNWMAATHACVPTDPNPOEIVLENT 84
Qy 61 EHFMMKNNWVEQWEDIIISLMDOSLKPCKVLTPLCTVLNCKDV-NATNTNDSGEMER 119
Db 85 EHFMMKNNWVEQWEDIIISLMDOSLKPCKVLTPLCTVLNCKDV-NATNTNDSGEMER 144
Qy 120 GEINCSFNITTSIRDEVOKYALFYKLDVYXIDNNNTSYRLISCDTSVITQACPKISFE 179
Db 145 GEINCSFKTTSIRNMOKYALFYKLDVPIIDNDTSYKLNCTNSVITQACPKISFE 204
Qy 180 PIPHYCAPAGFALIKNDKTFNGKPKCKNVSTQCTHGIRPVVSTOLLNGSLAEEBV 239
Db 205 PIPHYCAPAGFALIKNDKTFNGKPKCKNVSTQCTHGIRPVVSTOLLNGSLAEEBV 264
Qy 240 IRSNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRQAHN 281
Db 265 IRSNFTNNAKTIIVOLKESVEINCTRPNNNTRKSITIGPRAFYATGDIIGDIRQAHN 324
Qy 282 ISRAKNDTLKQIYIKLREOFENKTIYVNHSSGGDEIYVHSHFNGGGEFFYCNSTOLFNS 341
Db 325 ISGERKNNTLKQIYIKLQAFGNKTIYVKSOGSDPEIYVHSHFNGGGEFFYCNSTOLFNS 384
Qy 342 TWNNTEGSNNTEGNTITLPCRKOIINMOEVEKAMYPPIRQOIRCSSNITGLLTRD 401
Db 385 TW-NNTIGPNNTNG-TITLPCRKOIINMOEVEKAMYPPIRQOIRCSSNITGLLTRD 442
Qy 402 GGIN-ENGTEIFRPGGDMRDNRSELYKYVVIKIEPLGVAPTYKCKRRVYOREKAVGIG 460
Db 443 GGEKISNTTEIFRPGGDMRDNRSELYKYVVIKIEPLGVAPTYKCKRRVYOREKAVGIG 502
Qy 461 AVFLGFLGAGSTGASMTLTVQARLLSGIVQOONNLRAIEAOQMLQLTWGIKOL 520
Db 503 AMFLGFLGAGSTGASMTLTVQARLLSGIVQOONNLRAIEAOQMLQLTWGIKOL 562
Qy 521 QARVLAVERYLGDQOLLGIWCGSKLICTTAVPNNASMSNKSIDRINNNMTMEMEREID 580
Db 563 QARVLAVERYLKDQOLLGIWCGSKLICTTAVPNNASMSNKSIDRINNNMTMEMEREID 622
Qy 581 NYTSEIYTLIEESONQOEKNEOELLELDKMASLNNWPDITKMLY 625
Db 623 NYTMLIYTLIEESONQOEKNEOELLELDKMASLNNWPDITKMLY 667

```

RESULT 6

```

US-10-190-305A-2
; Sequence 2, Application US/10190305A
; Publication No. US20030198621A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEDEDE, Jan
; APPLICANT: BARNETT, Susan
; APPLICANT: LIAN, Ying
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR
; FILE REFERENCE: 2302-18702 / 18702.002
; CURRENT APPLICATION NUMBER: US/10/190.305A
; CURRENT FILING DATE: 2002-07-05
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 842
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SFI62
US-10-190-305A-2

```

Query Match 87.4%; Score 2928; DB 12; Length 842;
 Best Local Similarity 86.0%; Pred. No. 1.1e-270;
 Matches 555; Conservative 28; Mismatches 40; Indels 22; Gaps 5;

```

Qy 1 VEKLMVTYVYGVPMWKEATITLFCASDAKAYDTEVHNWMAATHACVPTDPNPOEIVLENT 60
Db 25 VEKLMVTYVYGVPMWKEATITLFCASDAKAYDTEVHNWMAATHACVPTDPNPOEIVLENT 84
Qy 61 EHFMMKNNWVEQWEDIIISLMDOSLKPCKVLTPLCTVLNCKDV-NATNTNDSGEMER 119
Db 85 EHFMMKNNWVEQWEDIIISLMDOSLKPCKVLTPLCTVLNCKDV-NATNTNDSGEMER 144
Qy 120 GEINCSFNITTSIRDEVOKYALFYKLDVYXIDNNNTSYRLISCDTSVITQACPKISFE 179
Db 145 GEINCSFKTTSIRNMOKYALFYKLDVPIIDNDTSYKLNCTNSVITQACPKISFE 204
Qy 180 PIPHYCAPAGFALIKNDKTFNGKPKCKNVSTQCTHGIRPVVSTOLLNGSLAEEBV 239
Db 205 PIPHYCAPAGFALIKNDKTFNGKPKCKNVSTQCTHGIRPVVSTOLLNGSLAEEBV 264
Qy 240 IRSNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRQAHN 281
Db 265 IRSNFTNNAKTIIVOLKESVEINCTRPNNNTRKSITIGPRAFYATGDIIGDIRQAHN 324
Qy 282 ISRAKNDTLKQIYIKLREOFENKTIYVNHSSGGDEIYVHSHFNGGGEFFYCNSTOLFNS 341
Db 325 ISGERKNNTLKQIYIKLQAFGNKTIYVKSOGSDPEIYVHSHFNGGGEFFYCNSTOLFNS 384
Qy 342 TWNNTEGSNNTEGNTITLPCRKOIINMOEVEKAMYPPIRQOIRCSSNITGLLTRD 401
Db 385 TW-NNTIGPNNTNG-TITLPCRKOIINMOEVEKAMYPPIRQOIRCSSNITGLLTRD 442
Qy 402 GGIN-ENGTEIFRPGGDMRDNRSELYKYVVIKIEPLGVAPTYKCKRRVYOREKAVGIG 460
Db 443 GGEKISNTTEIFRPGGDMRDNRSELYKYVVIKIEPLGVAPTYKCKRRVYOREKAVGIG 502
Qy 461 AVFLGFLGAGSTGASMTLTVQARLLSGIVQOONNLRAIEAOQMLQLTWGIKOL 520
Db 503 AMFLGFLGAGSTGASMTLTVQARLLSGIVQOONNLRAIEAOQMLQLTWGIKOL 562
Qy 521 QARVLAVERYLGDQOLLGIWCGSKLICTTAVPNNASMSNKSIDRINNNMTMEMEREID 580
Db 563 QARVLAVERYLKDQOLLGIWCGSKLICTTAVPNNASMSNKSIDRINNNMTMEMEREID 622
Qy 581 NYTSEIYTLIEESONQOEKNEOELLELDKMASLNNWPDITKMLY 625
Db 623 NYTMLIYTLIEESONQOEKNEOELLELDKMASLNNWPDITKMLY 667

```

RESULT 7

```

US-09-476-242-2
; Sequence 2, Application US/09476242
; Patent No. US20020146683A1
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: HARTOG, Karin
; APPLICANT: MARTIN, Eric
; TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES
; FILE REFERENCE: 1605.002
; CURRENT APPLICATION NUMBER: US/09/476.242
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 847
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-09-476-242-2

```

Query Match 87.4%; Score 2928; DB 10; Length 847;
 Best Local Similarity 86.0%; Pred. No. 1.1e-270;
 Matches 555; Conservative 26; Mismatches 40; Indels 22; Gaps 5;

```

Qy 61 EHFNNKNNNNVQWQWEDIIISLMDQSLKPCVKLTPLCTVLNCKDV-NATNTTNDSEGMERG 119
Db 90 EHFNNKNNNNVQWQWEDIIISLMDQSLKPCVKLTPLCTVLNCKDV-NATNTTNDSEGMERG 149
Qy 120 GGIKNSFNITTSIRDEVEKEVALFYKLDVVKIDNNNTSYRLISCDTSVITQACPKISFE 179
Db 150 GGIKNSFNITTSIRKQKQKVALFYKLDVVIDNDNTSYRLINCDTSVITQACPKISFE 209
Qy 180 PPIHYCAPAGFAILKCNKNTFNKGPCKNVSTXOCTHGIRPVVSTOLLNGLAEVYI 239
Db 210 PPIHYCAPAGFAILKCNKNTFNKGPCKNVSTXOCTHGIRPVVSTOLLNGLAEVYI 269
Qy 240 IRSDNFNNAKTIIYOLKESVEINCTRPNNNG-----AGDIRQAHCH 281
Db 270 IRSEFTDPAKTIIVOLKESVEINCTRPNNNTKRSITIGRPAFYATGDIIDIRQAHCH 329
Qy 282 ISRAKNDTLKOIVIKLRQFENKTIIVFNHSSGSGDEIYVHSHFNCGEFPYCNSTOLFNS 341
Db 330 ISGEKMNNTLKOIVIKLRQFENKTIIVFNHSSGSGDEIYVHSHFNCGEFPYCNSTOLFNS 389
Qy 342 TWNNTEGNNTEGNTITLPCRKOIINMMQWVGKAMVAPPIRGQIRCSSNTGLLITRD 401
Db 390 TW-NTTIGPNTNG-TITLPCRKOIINMMQWVGKAMVAPPIRGQIRCSSNTGLLITRD 447
Qy 402 GGIN-ENGTEIFRPGGDMRDNMRSELYKXKVKIIEPLGVAPTKCKRRVVOREKRAVGIG 460
Db 448 GGEISNTTEIFRPGGDMRDNMRSELYKXKVKIIEPLGVAPTKCKRRVVOREKRAVGIG 507
Qy 461 AVFLGFLGAAGSTMGASMTLTVOARLLSGIVQOQNNLLRAIEAQORMLQLTWVGIKOL 520
Db 508 AMFLGFLGAAGSTMGASMTLTVOARLLSGIVQOQNNLLRAIEAQORMLQLTWVGIKOL 567
Qy 521 QARVAVERYLDDOOLLGIMGSGKLICTTAVPMAWSNKSJLDRIWNNMTMEMEREID 580
Db 568 QARVAVERYLDDOOLLGIMGSGKLICTTAVPMAWSNKSJLDRIWNNMTMEMEREID 627
Qy 581 NYTSEIYTLIEESONQOEKNEQELLEDKWSLMMNFDITKWLWY 625
Db 628 NYTNIYTLIEESONQOEKNEQELLEDKWSLMMNFDITKWLWY 672

```

```

RESULT 8
US-10-032-162-15
; Sequence 15, Application US/10032162
; Publication No. US20030052839A1
; GENERAL INFORMATION:
; APPLICANT: BINGLEY, JAMES M
; APPLICANT: SCHUELEK, NORBERT
; APPLICANT: OLSON, WILLIAM C
; APPLICANT: PAUL, MADDON J
; APPLICANT: JOHN, MOORE P
; TITLE OF INVENTION: STABILIZED VIRAL ENVELOPE PROTEINS AND USES THEREOF
; FILE REFERENCE: 2048/59331az
; CURRENT APPLICATION NUMBER: US/10/032.162
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/602,864
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-10-032-162-15

```

```

Query Match 87.2%; Score 2923; DB 15; length 579;
Best Local Similarity 86.0%; Pred. No. 1.9e-270;
Matches 553; Conservative 0; Mismatches 8; Indels 82; Gaps 2;
Qy 1 VEKLVVVVVYGVVWKEATTTIFCASDAKAYTEVHNWATACVPTDNPQEVLENT 60
Db 1 VEKLVVVVVYGVVWKEATTTIFCASDAKAYTEVHNWATACVPTDNPQEVLENT 60

```

```

Qy 61 EHFNNKNNNNVQWQWEDIIISLMDQSLKPCVKLTPLCTVLNCKDV-NATNTTNDSEGMERG 120
Db 61 EHFNNKNNNNVQWQWEDIIISLMDQSLKPCVKLTPLCTVLNCKDV-NATNTTNDSEGMERG 98
Qy 121 EIKNSFNITTSIRDEVEKEVALFYKLDVVKIDNNNTSYRLISCDTSVITQACPKISFE 180
Db 99 -----GCDTSVITQACPKISFE 116
Qy 181 IPIHYCAPAGFAILKCNKNTFNKGPCKNVSTXOCTHGIRPVVSTOLLNGLAEVYI 240
Db 117 IPIHYCAPAGFAILKCNKNTFNKGPCKNVSTXOCTHGIRPVVSTOLLNGLAEVYI 176
Qy 241 RSDNFNNAKTIIYOLKESVEINCTRPNNNG-----AGDIRQAHCH 282
Db 177 RSDNFNNAKTIIYOLKESVEINCTRPNNNTKRSIHIGRPAFYATGDIIDIRQAHCH 236
Qy 283 SRAKNDTLKOIVIKLRQFENKTIIVFNHSSGSGDEIYVHSHFNCGEFPYCNSTOLFNS 342
Db 237 SRAKNDTLKOIVIKLRQFENKTIIVFNHSSGSGDEIYVHSHFNCGEFPYCNSTOLFNS 296
Qy 343 WNNTEGNNTEGNTITLPCRKOIINMMQWVGKAMVAPPIRGQIRCSSNTGLLITRD 402
Db 297 WNNTEGNNTEGNTITLPCRKOIINMMQWVGKAMVAPPIRGQIRCSSNTGLLITRD 356
Qy 403 GINENGTEIFRPGGDMRDNMRSELYKXKVKIIEPLGVAPTKCKRRVVOREKRAVGIGAV 462
Db 357 GINENGTEIFRPGGDMRDNMRSELYKXKVKIIEPLGVAPTKCKRRVVOREKRAVGIGAV 416
Qy 463 FLGFLGAAGSTMGASMTLTVOARLLSGIVQOQNNLLRAIEAQORMLQLTWVGIKOL 522
Db 417 FLGFLGAAGSTMGASMTLTVOARLLSGIVQOQNNLLRAIEAQORMLQLTWVGIKOL 476
Qy 523 RVLAVERLYLDDOOLLGIMGSGKLICTTAVPMAWSNKSJLDRIWNNMTMEMEREID 582
Db 477 RVLAVERLYLDDOOLLGIMGSGKLICTTAVPMAWSNKSJLDRIWNNMTMEMEREID 536
Qy 583 TSEIYTLIEESONQOEKNEQELLEDKWSLMMNFDITKWLWY 625
Db 537 TSEIYTLIEESONQOEKNEQELLEDKWSLMMNFDITKWLWY 579

```

```

RESULT 9
US-10-026-741-103
; Sequence 103, Application US/10026741
; Publication No. US20030049604A1
; GENERAL INFORMATION:
; APPLICANT: CHARNEAU, PIERRE
; APPLICANT: CLAVEL, FRANCOISE
; APPLICANT: BORMAN, ANDREW
; APPLICANT: OULLENT, CAROLINE
; APPLICANT: GUETARD, DENISE
; APPLICANT: MONTGANTER, LUC
; APPLICANT: COHEN, JACQUES
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF HIV-1 TYPE (OR SUBTYPE) ANTIGENS
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; Dunner, L.L.P.
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; APPLICATION NUMBER: US/10/026.741
; FILING DATE: 27-Dec-2001

```


PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/817,441
 FILING DATE: 31-AUG-1998
 APPLICATION NUMBER: PCT/FR 95/01391
 FILING DATE: 20-OCT-1995
 APPLICATION NUMBER: FR 9412554
 FILING DATE: 20-OCT-1994
 APPLICATION NUMBER: FR 9502526
 FILING DATE: 03-MAR-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Meyers, Kenneth J.
 REGISTRATION NUMBER: 25,146
 REFERENCE/DOCKET NUMBER: 03260, 6005-00000
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-408-4000
 TELEFAX: 202-408-4400
 INFORMATION FOR SEQ ID NO: 103:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 861 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 103:
 US-10-026-741-103

Query Match 86.6%; Score 2901.5; DB 15; Length 861;
 Best Local Similarity 84.3%; Pred. No. 3.9e-268; Indels 31; Gaps 7;
 Matches 552; Conservative 31; Mismatches 41;

2 EKLWVTVYGVVWKEATTTLFCASDAKAYDTEVHNVATHACVPTDNPQEVLEAVTE 61
 32 EKLWVTVYGVVWKEATTTLFCASDAKAYDTEVHNVATHACVPTDNPQEVLEAVTE 91
 62 HFNWKNWVWQWQEDIIISLMDQSLKPCVKLTPLCVTLNCKDVATNT---TNDSEG- 115
 92 NFNWKNWQWQWQEDIIISLMDQSLKPCVKLTPLCVTLNCKDVATNTSSSGE 151
 116 -TMRGEIKNGSPNITTSIRDEVQKAYLFFKLDVYXIDNNNTSRYLSGCTSVITQACP 174
 152 MMEKGEIKNGSPNITTSIRDEVQKAYLFFKLDVYXIDNNNTSRYLSGCTSVITQACP 211
 175 KISFEPIPIHYCAPAGFALILKCNKDTFNGKXPCKNVSTXOCHGIRPVYSTOLLNGSLA 234
 212 KVSFEPIPIHYCAPAGFALILKCNKDTFNGKXPCKNVSTXOCHGIRPVYSTOLLNGSLA 271
 235 EEEVIRSDNTNNAKTIIVOLKESVEINCTRPNNN-----GAGDI 275
 272 EEEVIRSDNTNNAKTIIVOLKESVEINCTRPNNN-----GAGDI 331
 276 RQAHCNISRAKNNPDLKQIVILKREOF-ENKTIIVFNHSGGDPPEIVMHSFNGGGEFFYCN 334
 332 RQAHCNISRAKNNPDLKQIVILKREOF-ENKTIIVFNHSGGDPPEIVMHSFNGGGEFFYCN 391
 335 STQFNSTWNN---NTEGSNNTEG-NTITLPCRILKQIINMOWEYKAMYPPIRQIRCS 390
 392 STQFNSTWNN---NTEGSNNTEG-NTITLPCRILKQIINMOWEYKAMYPPIRQIRCS 451
 391 SNITGLLTRDGGINENGEIIRPGGDMRDWRSSELYKYVVKIEPLGVAFTCKRRVY 450
 452 SNITGLLTRDGGINENGEIIRPGGDMRDWRSSELYKYVVKIEPLGVAFTCKRRVY 511
 451 QREKRAVIGAVPLGFLGAAGSTMGASMTLTVQARLLLSGIVOOONNLLRAIEOQHLL 510
 512 QREKRAVIGAVPLGFLGAAGSTMGASMTLTVQARLLLSGIVOOONNLLRAIEOQHLL 571
 511 QLTWVIGIQLQARVLAVERYLADQOLGIGWCSGLICTTAVPMNASNKSLEDIRIMNM 570
 572 QLTWVIGIQLQARVLAVERYLADQOLGIGWCSGLICTTAVPMNASNKSLEDIRIMNM 631
 571 TWMEEREIDNTYSEIYTLIEESONQOEKNEOELLELDKMASLWNPEDITKWLWY 625
 632 TWMEEREIDNTYSEIYTLIEESONQOEKNEOELLELDKMASLWNPEDITKWLWY 686

RESULT 10
 US-09-476-242-1
 Sequence 1, Application US/09476242
 Patent No. US20020146683A1
 GENERAL INFORMATION:
 APPLICANT: BARNETT, Susan
 APPLICANT: HARTOG, Karin
 APPLICANT: MARTIN, Eric
 TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES
 FILE REFERENCE: 1605,002
 CURRENT APPLICATION NUMBER: US/09/476,242
 CURRENT FILING DATE: 1999-12-30
 NUMBER OF SEQ ID NOS: 26
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 1
 LENGTH: 856
 TYPE: PRT
 ORGANISM: Human immunodeficiency virus
 US-09-476-242-1

Query Match 85.9%; Score 2878; DB 10; Length 856;
 Best Local Similarity 83.7%; Pred. No. 6.8e-266; Indels 26; Gaps 5;
 Matches 544; Conservative 35; Mismatches 45;

2 EKLWVTVYGVVWKEATTTLFCASDAKAYDTEVHNVATHACVPTDNPQEVLEAVTE 61
 32 EKLWVTVYGVVWKEATTTLFCASDAKAYDTEVHNVATHACVPTDNPQEVLEAVTE 91
 62 HFNWKNWVWQWQEDIIISLMDQSLKPCVKLTPLCVTLNCKDVATNTNDSEG--TMR 119
 92 NFNWKNWQWQWQEDIIISLMDQSLKPCVKLTPLCVTLNCKDVATNTSSSGRIMEK 151
 120 GEIKNGSPNITTSIRDEVQKAYLFFKLDVYXIDNNNTSRYLSGCTSVITQACPVSFE 179
 152 GEIKNGSPNITTSIRDEVQKAYLFFKLDVYXIDNNNTSRYLSGCTSVITQACPVSFE 211
 180 PIPHYCAPAGFALILKCNKDTFNGKXPCKNVSTXOCHGIRPVYSTOLLNGSLAEEV 239
 212 PIPHYCAPAGFALILKCNKDTFNGKXPCKNVSTXOCHGIRPVYSTOLLNGSLAEEV 271
 240 IRSNFTNNAKTIIVOLKESVEINCTRPNNN-----GAGDIRQAH 280
 272 IRSNFTNNAKTIIVOLKESVEINCTRPNNN-----GAGDIRQAH 331
 281 NISRAKNNPDLKQIVILKREOF-ENKTIIVFNHSGGDPPEIVMHSFNGGGEFFYCNSTOLF 339
 332 NISRAKNNPDLKQIVILKREOF-ENKTIIVFNHSGGDPPEIVMHSFNGGGEFFYCNSTOLF 391
 340 NSTWNN---NTEGSNNTEG-NTITLPCRILKQIINMOWEYKAMYPPIRQIRCSNITG 395
 392 NSTWNN---NTEGSNNTEG-NTITLPCRILKQIINMOWEYKAMYPPIRQIRCSNITG 451
 396 LLLTRDGGINENGEIIRPGGDMRDWRSSELYKYVVKIEPLGVAFTCKRRVQEK 455
 452 LLLTRDGGINENGEIIRPGGDMRDWRSSELYKYVVKIEPLGVAFTCKRRVQEK 511
 456 AVIGAVPLGFLGAAGSTMGASMTLTVQARLLLSGIVOOONNLLRAIEOQHLL 515
 512 AVIGAVPLGFLGAAGSTMGASMTLTVQARLLLSGIVOOONNLLRAIEOQHLL 571
 516 GIKQOLQARVLAVERYLADQOLGIGWCSGLICTTAVPMNASNKSLEDIRIMNM 575
 572 GIKQOLQARVLAVERYLADQOLGIGWCSGLICTTAVPMNASNKSLEDIRIMNM 631
 576 EREIDNTYSEIYTLIEESONQOEKNEOELLELDKMASLWNPEDITKWLWY 625
 632 EREIDNTYSEIYTLIEESONQOEKNEOELLELDKMASLWNPEDITKWLWY 681

RESULT 11
 US-10-196-515-11

```

; Sequence 11, Application US/10196515
; Publication No. US20030091594A1
; GENERAL INFORMATION:
; APPLICANT: HOXIE, James A.
; APPLICANT: LABRANCHE, Celja C.
; APPLICANT: DOMS, Robert W.
; APPLICANT: HOFFMAN, Trevor L.
; TITLE OF INVENTION: CD4-INDEPENDENT HIV ENVELOPE PROTEINS AS VACCINES AND
; FILE REFERENCE: HOXIE 9596-104U1 (0282)
; CURRENT APPLICATION NUMBER: US/10/196,515
; PRIOR FILING DATE: 1999-05-24
; PRIOR APPLICATION NUMBER: US/09/337,387
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: US/09/317,556
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 856
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-10-196-515-11

```

```

Query Match      85.6%; Score 2867; DB 15; Length 856;
Beet Local Similarity 83.5%; Pred. No. 7,7e-265;
Matches 543; Conservative 35; Mismatches 46; Indels 26; Gaps 5;

QY 2 EKLAWTVYGVVWKEATTTTFCASDAKAYDEVHNVNATHACVPTDPNPQEVLENYTE 61
DB 32 EKLAWTVYGVVWKEATTTTFCASDAKAYDEVHNVNATHACVPTDPNPQEVLENYTE 91
QY 62 HNNMKNNVMOEDIIISLDQSLKPCVKLTPLCVTLCKDVNATNTNDSG--TWER 119
DB 92 NFDMMKNDVMOEDIIISLDQSLKPCVKLTPLCVSLCKDTLCKDNTNNSGGMINEK 151
QY 120 GEIKKCSFNITTSIRDEVOKEYALFYKLDVXVINDNNNTSYLLISCDTSVITQACPKISFE 179
DB 152 GEIKKCSFNITTSIRDEVOKEYALFYKLDIIPIDNPTSYLLISCDTSVITQACPKISFE 211
QY 180 PIPHYCAPAGFAIIKCKNDKTFNGCKPCKNSTXOCTGIRPVVSTOLLNGLSLAEEVY 239
DB 212 PIPHYCAPAGFAIIKCKNDKTFNGCKPCKNSTXOCTGIRPVVSTOLLNGLSLAEEVY 271
QY 240 IRSDFTNNAKTIIVOLKESVEINCTRPNN-----GAGDIRQAHG 280
DB 272 IRSDFTNNAKTIIVOLKESVEINCTRPNN-----GAGDIRQAHG 331
QY 281 NISRAKNDTLKQIYIKLREOF-ENKTIIVFNHSSGDEBEIYVHNSFCGGEFFYCNSTOLF 339
DB 332 NISRAKNDTLKQIYIKLREOF-ENKTIIVFNHSSGDEBEIYVHNSFCGGEFFYCNSTOLF 391
QY 340 NSTWNN--NTEGSNNTEG-NTITLPCRIKQIIMMOEYVGAAPYAPRIGQIRCSNITG 395
DB 392 NSTWNN--NTEGSNNTEG-NTITLPCRIKQIIMMOEYVGAAPYAPRIGQIRCSNITG 451
QY 396 LLTFDGGINENGTIFRPGGDMRDNRSELVYKVKVIEPLGVAAPTKARRVVOREKRA 455
DB 452 LLTFDGGINENGTIFRPGGDMRDNRSELVYKVKVIEPLGVAAPTKARRVVOREKRA 511
QY 456 AVGIGAVELGFLGAAGSTMGAASTLTVQARLLLSGIYQOQNNLLRAIEAOQHLQLTVM 515
DB 512 AVGIGAVELGFLGAAGSTMGAASTLTVQARLLLSGIYQOQNNLLRAIEAOQHLQLTVM 571
QY 516 GIKQIQAFLAVERYLGDQOLIGWCSGKLICTTAVPANASWNSKSLDRIMNNMTWME 575
DB 572 GIKQIQAFLAVERYLGDQOLIGWCSGKLICTTAVPANASWNSKSLDRIMNNMTWME 631
QY 576 BREIDNTYSEIYTLIEESONOEKNEOELLELDKWSIIMNPNITNMLWY 625
DB 632 BREIDNTYSEIYTLIEESONOEKNEOELLELDKWSIIMNPNITNMLWY 681

```

```

RESULT 12
US-10-196-515-3
; Sequence 3, Application US/10196515
; Publication No. US20030091594A1
; GENERAL INFORMATION:
; APPLICANT: HOXIE, James A.
; APPLICANT: LABRANCHE, Celja C.
; APPLICANT: DOMS, Robert W.
; APPLICANT: HOFFMAN, Trevor L.
; TITLE OF INVENTION: CD4-INDEPENDENT HIV ENVELOPE PROTEINS AS VACCINES AND
; FILE REFERENCE: HOXIE 9596-104U1 (0282)
; CURRENT APPLICATION NUMBER: US/10/196,515
; PRIOR FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: US/09/337,387
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: US/09/317,556
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 726
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-10-196-515-3

```

```

Query Match      84.2%; Score 2822.5; DB 15; Length 726;
Beet Local Similarity 82.4%; Pred. No. 1.1e-260;
Matches 533; Conservative 38; Mismatches 51; Indels 25; Gaps 5;

QY 2 EKLAWTVYGVVWKEATTTTFCASDAKAYDEVHNVNATHACVPTDPNPQEVLENYTE 61
DB 32 EKLAWTVYGVVWKEATTTTFCASDAKAYDEVHNVNATHACVPTDPNPQEVLENYTE 91
QY 62 HNNMKNNVMOEDIIISLDQSLKPCVKLTPLCVTLCKDVNATNTNDSG--TWER 119
DB 92 NFDMMKNDVMOEDIIISLDQSLKPCVKLTPLCVSLCKDTLCKDNTNNSGGMINEK 151
QY 120 GEIKKCSFNITTSIRDEVOKEYALFYKLDVXVINDNNNTSYLLISCDTSVITQACPKISFE 179
DB 152 GEIKKCSFNITTSIRDEVOKEYALFYKLDIIPIDNPTSYLLISCDTSVITQACPKISFE 211
QY 180 PIPHYCAPAGFAIIKCKNDKTFNGCKPCKNSTXOCTGIRPVVSTOLLNGLSLAEEVY 239
DB 212 PIPHYCAPAGFAIIKCKNDKTFNGCKPCKNSTXOCTGIRPVVSTOLLNGLSLAEEVY 271
QY 240 IRSDFTNNAKTIIVOLKESVEINCTRPNN-----GAGDIRQAHG 280
DB 272 IRSDFTNNAKTIIVOLKESVEINCTRPNN-----GAGDIRQAHG 331
QY 281 NISRAKNDTLKQIYIKLREOF-ENKTIIVFNHSSGDEBEIYVHNSFCGGEFFYCNSTOLF 339
DB 332 NISRAKNDTLKQIYIKLREOF-ENKTIIVFNHSSGDEBEIYVHNSFCGGEFFYCNSTOLF 391
QY 340 NSTWNN--NTEGSNNTEG-NTITLPCRIKQIIMMOEYVGAAPYAPRIGQIRCSNITG 398
DB 392 NSTWNN--NTEGSNNTEG-NTITLPCRIKQIIMMOEYVGAAPYAPRIGQIRCSNITG 449
QY 396 LLTFDGGINENGTIFRPGGDMRDNRSELVYKVKVIEPLGVAAPTKARRVVOREKRA 458
DB 452 LLTFDGGINENGTIFRPGGDMRDNRSELVYKVKVIEPLGVAAPTKARRVVOREKRA 509
QY 456 AVGIGAVELGFLGAAGSTMGAASTLTVQARLLLSGIYQOQNNLLRAIEAOQHLQLTVM 518
DB 512 AVGIGAVELGFLGAAGSTMGAASTLTVQARLLLSGIYQOQNNLLRAIEAOQHLQLTVM 579
QY 516 GIKQIQAFLAVERYLGDQOLIGWCSGKLICTTAVPANASWNSKSLDRIMNNMTWME 578
DB 572 GIKQIQAFLAVERYLGDQOLIGWCSGKLICTTAVPANASWNSKSLDRIMNNMTWME 629
QY 576 BREIDNTYSEIYTLIEESONOEKNEOELLELDKWSIIMNPNITNMLWY 625
DB 632 BREIDNTYSEIYTLIEESONOEKNEOELLELDKWSIIMNPNITNMLWY 676

```

RESULT 13
US-10-196-515-12
Sequence 12, Application US/10196515
Publication No. US20030091594A1
GENERAL INFORMATION:
APPLICANT: HOKIE, James A.
APPLICANT: LABRANCHE, Celie C.
APPLICANT: DOMS, Robert W.
APPLICANT: HOFFMAN, Trevor L.
TITLE OF INVENTION: CD4-INDEPENDENT HIV ENVELOPE PROTEINS AS VACCINES AND
TITLE OF INVENTION: THERAPEUTICS
FILE REFERENCE: Hoxie 9596-104U1 (0282)
CURRENT APPLICATION NUMBER: US/10/196,515
CURRENT FILING DATE: 2002-07-16
PRIOR APPLICATION NUMBER: US/09/337,387
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: US 09/317,556
PRIOR FILING DATE: 1999-05-24
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
LENGTH: 759
TYPE: PRT
ORGANISM: Human immunodeficiency virus type 1
US-10-196-515-12

Query Match 84.0%; Score 2815.5; DB 15; Length 759;
Best Local Similarity 82.4%; Pred. No. 5.4e-260;
Matches 533; Conservative 38; Mismatches 51; Indels 25; Gaps 5;

QY 2 EKLWTVYGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDPNPQVLENTHEF 61
DB 32 EKLWTVYGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDPNPQVLENTHEF 91
QY 62 HNNWKNWVQWQEDIIISLWQSLKPCVKLTPLCVTLNCKDVANNTTNDSEG--TWER 119
DB 92 HNNWKNWVQWQEDIIISLWQSLKPCVKLTPLCVTLNCKDVANNTTNDSEG--TWER 151
QY 120 GEINQCSFNITTSIRDEYQKRYALFYKLDVYXIDNNNTSYLISCDTSVITQACPKEISFE 179
DB 152 GEINQCSFNITTSIRDEYQKRYALFYKLDVYXIDNNNTSYLISCDTSVITQACPKEISFE 211
QY 180 PIPHYCAPAGFALIKKNDKTEGKXPKGNVSTXQCTHGRIPVYSTOLLNGSLAEEBHV 239
DB 212 PIPHYCAPAGFALIKKNDKTEGKXPKGNVSTXQCTHGRIPVYSTOLLNGSLAEEBHV 271
QY 240 IRSNFTNNAKTIIVQLKESVEINCTRENN-----GAGDIRQAH 280
DB 272 IRSNFTNNAKTIIVQLKESVEINCTRENN-----GAGDIRQAH 331
QY 281 NISAKKNDTKQIVIKREOF-ENKTVFNHSGGDBEIMHSGGDEFFYCNSTOLF 339
DB 332 NISAKKNDTKQIVIKREOF-ENKTVFNHSGGDBEIMHSGGDEFFYCNSTOLF 391
QY 340 NSTNNNTGSGNTEG--NTITLPCRKOIIMMOEYKAWAPPIRGQIRSSNITGILL 398
DB 392 NST--STKSGNNTGSGNTEG--NTITLPCRKOIIMMOEYKAWAPPIRGQIRSSNITGILL 449
QY 399 TRDGINENGTETFRPGGDMRDNRSELVYKVKIIBPLGVAPTKRRVVOBEKRAVG 458
DB 450 TRDGINENGTETFRPGGDMRDNRSELVYKVKIIBPLGVAPTKRRVVOBEKRAVG 509
QY 459 IGAVFLGFLGAAGSTMGAASTLTVQARLLISGIVQOONNLLRAIEAOHMLQLTWGIK 518
DB 510 IGAVFLGFLGAAGSTMGAASTLTVQARLLISGIVQOONNLLRAIEAOHMLQLTWGIK 569
QY 519 QLOARVLAVERYLGDQQLIGTWGSGKLICTTAVPWNASWSKSLDRIIMNMWMEWRE 578
DB 570 QLOARVLAVERYLGDQQLIGTWGSGKLICTTAVPWNASWSKSLDRIIMNMWMEWRE 629
QY 579 IDNTSEIYTLIESONQOEKNEBELLELDKMASIMNFDITKMLWY 625

DB 630 INNTSLHSLIESQWQEMNEBELLELDKMASIMNFDITKMLWY 676

RESULT 14
US-09-938-406-1
Sequence 1, Application US/09938406
Patent No. US20020155120A1
GENERAL INFORMATION:
APPLICANT: Vancolt, Thomas
APPLICANT: Birtx, Deborah
TITLE OF INVENTION: PROTEIN AND PEPTIDE VACCINES FOR
INDUCING MUCOSAL IMMUNITY
FILE REFERENCE: 40646-20002.10
CURRENT APPLICATION NUMBER: US/09/938,406
CURRENT FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: US 09/214,701
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/US 97/12253
PRIOR FILING DATE: 1997-07-10
PRIOR APPLICATION NUMBER: US 60/021,687
PRIOR FILING DATE: 1996-07-10
NUMBER OF SEQ ID NOS: 18
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 868
TYPE: PRT
ORGANISM: Virus HIV-1
US-09-938-406-1

Query Match 83.4%; Score 2793.5; DB 10; Length 868;
Best Local Similarity 80.1%; Pred. No. 8.4e-258;
Matches 528; Conservative 36; Mismatches 58; Indels 37; Gaps 5;

QY 4 LMTVTVYGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDPNPQVLENTHEF 63
DB 35 LMTVTVYGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDPNPQVLENTHEF 94
QY 64 HNNWKNWVQWQEDIIISLWQSLKPCVKLTPLCVTLNCKDVANNTTNDSEGTM----- 117
DB 95 HNNWKNWVQWQEDIIISLWQSLKPCVKLTPLCVTLNCKDVANNTTNDSEGTM----- 154
QY 118 -ERGEINQCSFNITTSIRDEYQKRYALFYKLDVYXIDN-----NNTSYLISCDTSVITQ 171
DB 155 -ERGEINQCSFNITTSIRDEYQKRYALFYKLDVYXIDN-----NNTSYLISCDTSVITQ 214
QY 172 ACPKISEPPIPIHYCAPAGFALIKKNDKTEGKXPKGNVSTXQCTHGRIPVYSTOLLNG 231
DB 215 ACPKISEPPIPIHYCAPAGFALIKKNDKTEGKXPKGNVSTXQCTHGRIPVYSTOLLNG 274
QY 232 SLAEEVYVTSDFNTNNAKTIIVQLKESVEINCTRENN-----AG 273
DB 275 SLAEEVYVTSDFNTNNAKTIIVQLKESVEINCTRENN-----AG 334
QY 274 DIRQAHGNISRAKNDTKQIVIKREOFENKTVFNHSGGDBEIMHSGGDEFFYCNSTOLF 333
DB 335 DIRQAHGNISRAKNDTKQIVIKREOFENKTVFNHSGGDBEIMHSGGDEFFYCNSTOLF 394
QY 334 NSTQLFNSTNNNTGSG-----NNTGNTITLPCRKOIIMMOEYKAWAPPIRGQI 387
DB 395 NSTQLFNSTNNNTGSG-----NNTGNTITLPCRKOIIMMOEYKAWAPPIRGQI 454
QY 388 RCSNITGILLTRDGINENGTETFRPGGDMRDNRSELVYKVKIIBPLGVAPTKRR 447
DB 455 RCSNITGILLTRDGINENGTETFRPGGDMRDNRSELVYKVKIIBPLGVAPTKRR 514
QY 448 RVVQBEKRAVG-IGAVFLGFLGAAGSTMGAASTLTVQARLLISGIVQOONNLLRAIEAO 506
DB 515 RVVQBEKRAVG-IGAVFLGFLGAAGSTMGAASTLTVQARLLISGIVQOONNLLRAIEAO 574
QY 507 QRMQLTGWGIKQLOARVLAVERYLGDQQLIGTWGSGKLICTTAVPWNASWSKSLDRI 566

Db 575 QHLLDTWGCIGLOARILAVERYLKDQQLGFWGSGKLICTTAVPWNASNSKTLDOI 634
 QY 567 WNNMTWMEWEREIDNTSEIYTLIEESONQOEKNEQELLEDKWASLWNPDTTKWLY 625
 Db 635 WNNMTWMEWRIDNTYTHIYTLIEESONQOEKNEQELLEDKWASLWNPDTTKWLY 633

Db 573 NYTNLIYTLIEESONQOEKNEQELLEDKWASLWNPDTTKWLY 617
 Search completed: December 12, 2003, 12:45:24
 Job time : 27.0558 secs

RESULT 15

US-09-891-609-4
 ; Sequence 4, Application US/09891609
 ; Patent No. US20020127238A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Stamatacos, Leonidas
 ; TITLE OF INVENTION: HIV-1 VACCINES AND SCREENING METHODS THEREFOR
 ; FILE REFERENCE: 2570-1-002N
 ; CURRENT APPLICATION NUMBER: US/09/891,609
 ; PRIOR FILING DATE: 2001-06-26
 ; PRIOR APPLICATION NUMBER: 60/214,608
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 4
 ; LENGTH: 619
 ; TYPE: PRT
 ; ORGANISM: Human immunodeficiency virus type 1
 US-09-891-609-4

Query Match 82.3%; Score 2758; DB 10; Length 619;
 Best Local Similarity 82.0%; Pred. No. 1.3e-254;
 Matches 529; Conservative 24; Mismatches 42; Indels 50; Gaps 7;

QY 1 VEKLTWTVYGVVWVKEATTTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNVT 60
 Db 3 VEKLTWTVYGVVWVKEATTTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNVT 62
 QY 61 EHFNMKNMVEQWQEDIIISLWDSIKPCVKLTPLCVTLNCKDV-NATNTNDSGTMER 119
 Db 63 EHFNMKNMVEQWQEDIIISLWDSIKPCVKLTPLCVTLNCKDV-NATNTNDSGTMER 122
 QY 120 GEIKNCSFIIITSINDEVOKEALFYKLDVVAIDNNNTSYRLISCDTSYITQACPKISFE 179
 Db 123 GEIKNCSFVVGAG-----KLNCTSVITQACPKVSFE 155
 QY 180 PPIHYCAPAGFAILKNDKTFNGKXPCQNVSTXOCTHGIRPVVSTOLLNGSLAEEVY 239
 Db 156 PPIHYCAPAGFAILKNDKTFNGSGPCTNVSTVOCTHGIRPVVSTOLLNGSLAEEVY 215
 QY 240 IRSDFNTNAKTIIVOLKESVEINCTRPNNNG-----AGDIRQAHGN 281
 Db 216 IRSENFDTAKTIIVOLKESVEINCTRPNNNRKSTIGPRAFYATGDIIGDIRQAHGN 275
 QY 282 ISRAKNDTLKQIVIKLREQENKTIIVFNHSSGGDEIYMHSPNCGEFPYCNSTOLFNS 341
 Db 276 ISGEKKNNTLKOIVTKLQAFQNKTIIVFKOSSGGDEIYMHSPNCGEFPYCNSTOLFNS 335
 QY 342 TWNNNTGSGNTEGNTITLPCRIKQIINWQEVGKAMVAPPIRGQIRCSNITGLLTRD 401
 Db 336 TW-NTTIGPNTNG--TITLPCRIKQIINWQEVGKAMVAPPIRGQIRCSNITGLLTRD 393
 QY 402 GGIN-ENGTEIFRPGGDMRDNRSELVYKYVKYKIEPLGAVPTKCRRVVQREKRAVGIG 460
 Db 394 GGEKISNTTEIFRPGGDMRDNRSELVYKYVKYKIEPLGAVPTKCRRVVQREKRAVTLG 453
 QY 461 AVFLGFLGAAGSTGAASMTLTVQARLLISGIVQOQNNILRAIEAQRMLOLTWVGIKOL 520
 Db 454 AMFLGFLGAAGSTGAASMTLTVQARLLISGIVQOQNNILRAIEAQRMLOLTWVGIK-L 512
 QY 521 QARVLAVERYLGDQQLGFWGSGKLICTTAVPWNASNSKSLDRIANNMTWMEWERID 580
 Db 513 QARVLAVERYLGDQQLGFWGSGKLICTTAVPWNASNSKSLDRIANNMTWMEWERID 572
 QY 581 NYTSEIYTLIEESONQOEKNEQELLEDKWASLWNPDTTKWLY 625

THIS PAGE BLANK (US:TD)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2003, 12:25:38 / Search time 10.49 Seconds

(without alignments)
2801,881 Million cell updates/sec

Title: US-10-032-162-17

Perfect score: 3351

Sequence: 1 VEKMWTVYGVVWKEATT.....ELDKWASIMNWDITKIMLY 625

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues

1 number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3008.5	89.8	848	1	ENV_HV1JR
2	2928	87.4	847	1	ENV_HV1SI
3	2903.5	86.6	856	1	ENV_HV1SC
4	2901.5	86.6	861	1	ENV_HV1BR
5	2887	86.2	856	1	ENV_HV1BI
6	2884	86.1	867	1	ENV_HV1J3
7	2883	86.0	856	1	ENV_HV1PV
8	2882	86.0	843	1	ENV_HV1Y2
9	2878	85.9	852	1	ENV_HV1BN
10	2878	85.9	856	1	ENV_HV1H2
11	2871	85.7	856	1	ENV_HV1H3
12	2870	85.6	856	1	ENV_HV1MN
13	2864	85.5	856	1	ENV_HV1LW
14	2861.5	85.4	851	1	ENV_HV1B8
15	2861.5	85.4	856	1	ENV_HV1A3
16	2845.5	84.9	852	1	ENV_HV1S3
17	2836	84.6	855	1	ENV_HV1A2
18	2830	84.5	853	1	ENV_HV1MP
19	2818	84.1	847	1	ENV_HV1W2
20	2817	84.1	855	1	ENV_HV1YR
21	2810	83.9	855	1	ENV_HV1OY
22	2793.5	83.4	868	1	ENV_HV1C4
23	2790.5	83.3	861	1	ENV_HV1XB
24	2645	78.9	855	1	ENV_HV1Z6
25	2641	78.8	863	1	ENV_HV1Z8
26	2636	78.7	853	1	ENV_HV1Z2
27	2578.5	76.9	846	1	ENV_HV1ND
28	2572.5	76.8	859	1	ENV_HV1MA
29	2562	76.5	853	1	ENV_HV1Z4
30	2517.5	75.1	856	1	ENV_HV1ZB
31	2180	65.1	854	1	ENV_S1VC2
32	1631.5	48.7	421	1	ENV_HV1N5
33	1542	46.0	460	1	ENV_HV1Z3

34	1199.5	35.8	854	1	ENV_S1VA1	Q02837 simian immu
35	1195.5	35.7	851	1	ENV_HV2G1	P18040 human immu
36	1194	35.6	859	1	ENV_HV2ST	P20872 human immu
37	1176	35.1	712	1	ENV_HV2S2	P32536 human immu
38	1173.5	35.0	768	1	ENV_S1VA1	P27757 simian immu
39	1165	34.8	877	1	ENV_S1VAG	P27977 simian immu
40	1164	34.7	858	1	ENV_HV2RO	P04577 human immu
41	1158	34.6	865	1	ENV_S1VAT	P05886 simian immu
42	1148.5	34.3	859	1	ENV_HV2CA	P24105 human immu
43	1143	34.1	882	1	ENV_HV2NZ	P05883 human immu
44	1136	33.9	886	1	ENV_S1VW1	P05885 simian immu
45	1135.5	33.9	851	1	ENV_HV2D1	P17755 human immu

ALIGNMENTS

RESULT 1	ID	ENV_HV1JR	STANDARD;	PRT;	848 AA.
AC	P20871;				
DT	01-FEB-1991 (Rel. 17, Created)				
DT	01-FEB-1991 (Rel. 17, Last sequence update)				
DT	15-SEP-2003 (Rel. 42, Last annotation update)				
DE	Envelope glycoprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].				
GN	ENV				
OS	Human immunodeficiency virus type 1 (JRCSF isolate) (HIV-1).				
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.				
OX	NCBI_TaxID=11688;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Koyanagi S., Chen I.S.Y.;				
RL	Submitted (DEC-1988) to the HIV data bank.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; M38429; AAB03749.1; -				
DR	PDB; 1CE4; 16-MAR-99.				
DR	HIV; M38429; ENV5URCSF.				
DR	InterPro; IPR000328; Env_GP41.				
DR	InterPro; IPR000777; GP120.				
DR	Pfam; PF00516; GP120; 1.				
DR	Pfam; PF00517; GP41; 1.				
KW	AIDS; Coat protein; Glycoprotein; Polypeptide; Transmembrane; Signal;				
KW	3D-structure.				
FT	SIGNAL	1	32		
FT	CHAIN	33	503		EXTERIOR MEMBRANE GLYCOPROTEIN.
FT	CHAIN	504	848		TRANSMEMBRANE GLYCOPROTEIN.
FT	DISULFID	53	73		BY SIMILARITY.
FT	DISULFID	118	203		BY SIMILARITY.
FT	DISULFID	125	194		BY SIMILARITY.
FT	DISULFID	130	154		BY SIMILARITY.
FT	DISULFID	216	245		BY SIMILARITY.
FT	DISULFID	226	237		BY SIMILARITY.
FT	DISULFID	294	328		BY SIMILARITY.
FT	DISULFID	374	437		BY SIMILARITY.
FT	DISULFID	381	410		BY SIMILARITY.
FT	CARBOHYD	87	87		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	134	134		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	137	137		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	153	153		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	157	157		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	185	185		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	195	195		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	228	228		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	239	239		N-LINKED (GLCNAC. . .) (POTENTIAL).

```

FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 274 274 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 382 382 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 388 388 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 403 403 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 440 440 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 453 453 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 603 603 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 608 608 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 617 617 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 629 629 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 808 808 N-LINKED (GLCNAC. . .) (POTENTIAL).
SEQUENCE 848 AA; 96475 MW; 2076751227EC3F3 CRC64;

Query Match 89.8%; Score 3008.5; DB 1; Length 848;
Best Local Similarity 88.2%; Pred. No. 3.2e-232;
Matches 569; Conservative 22; Mismatches 33; Indels 21; Gaps 4;

QY 1 VEKLVTVYVYGVWPKKEATTLFCASDAKADTEVHNVMAATHACVPTDPNPOEVLENTV 60
DB 30 VEKLVTVYVYGVWPKKEATTLFCASDAKADTEVHNVMAATHACVPTDPNPOEVLENTV 89
QY 61 EHPMMKNNWEOQEDTISLMDOSLKPVCVLTPLCVTLNCKDVNATNTNDESGTMERG 120
DB 90 EDFMMKNNWEOQEDTISLMDOSLKPVCVLTPLCVTLNCKDVNATNTNDESGTMERG 149
QY 121 EIKKCSFNTISIDENCKEVALPKLDVXIDN-NNTSYLLIGDSVITQACPKISFE 179
DB 150 EIKKCSFNTISIDENCKEVALPKLDVXIDN-NNTSYLLIGDSVITQACPKISFE 209
QY 180 PIPHYCAPAGFALIKCKDFTFNKPKCKANVTXOCTGIRPVSTOLLNGSLAESENV 239
DB 210 PIPHYCAPAGFALIKCKDFTFNKPKCKANVTXOCTGIRPVSTOLLNGSLAESENV 269
QY 240 IRSNFTNNAKTIIIVOLKESVEINCTRPNNNG-----AGDIRQAHCN 281
DB 270 IRSNFTNNAKTIIIVOLKESVEINCTRPNNNG-----AGDIRQAHCN 329
QY 282 ISRAKNDTLLKOIYIKLEOFENKTIVFNHSSGGDEPEIVMHSFNGGFEFFCNSTOLFNS 341
DB 330 ISRAKNDTLLKOIYIKLEOFENKTIVFNHSSGGDEPEIVMHSFNGGFEFFCNSTOLFNS 389
QY 342 TWMNNTGEGSNNTGEG-TTLPCRIKQIINMNOEVKAYAPPIRQOJCSNITGLLLTR 400
DB 390 TWMNNTGEGSNNTGEG-TTLPCRIKQIINMNOEVKAYAPPIRQOJCSNITGLLLTR 448
QY 401 DGGINENGTETFRPGGDMRDNMWSELKYKVKIETPLGVAPTKCKRRVORERKAVGIG 460
DB 449 DGGINENGTETFRPGGDMRDNMWSELKYKVKIETPLGVAPTKCKRRVORERKAVGIG 508
QY 461 AVPLGFGAAGSTGASMTLTVQARLLSGIVOOQNNLLAIREQOQMLDLTWGIGIQL 520
DB 509 AVPLGFGAAGSTGASMTLTVQARLLSGIVOOQNNLLAIREQOQMLDLTWGIGIQL 568
QY 521 QARVLAVERYLGDQOOLGIMGCSGLICTAVPMNASNSKSLDRINNMNTMMEWEREID 580
DB 569 QARVLAVERYLGDQOOLGIMGCSGLICTAVPMNASNSKSLDRINNMNTMMEWEREID 628
QY 581 NYTSEIYTLIEESQNOQKNEQELLEDKMAKSLMNPFDITKWLWY 625
DB 629 NYTSEIYTLIEESQNOQKNEQELLEDKMAKSLMNPFDITKWLWY 673

```

```

AC P19550;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Envelope polypeptide in GPI10 precursor [contains: Exterior membrane
DE glycoprotein (GPI120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (SF162 isolate) (HIV-1).
OC Viruses; Retroviridae; Lentivirinae.
OX NCBI_TaxID=11691;
RX MEDLINE=90347835; PubMed=2384920;
RA Cheng-Mayer C., Quiroz M., Tung J.W., Levy J.
RT "Viral determinants of human immunodeficiency virus type 1 T-cell or
RT macrophage tropism: cytopathogenicity, and CD4 antigen modulation.";
RL J. Virol. 64:4390-4398(1990).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M65024; AAA45072.1; -.
CC PDB: 1OBS; 15-MAY-97.
CC DR HIV; M38428; ENV$SF162.
CC DR InterPro; IPR000328; Env_GP41.
CC DR InterPro; IPR000777; GP120.
CC DR Pfam; PF00516; GP120; 1.
CC DR Pfam; PF00517; GP41; 1.
CC AIDS; Coat protein; polypeptide; Glycoprotein; Transmembrane; Signal;
CC 3D-structure.
CC KX SIGNAL 1 29
CC FT CHAIN 30 502
CC FT CHAIN 503 847
CC FT DISULFID 53 73
CC FT DISULFID 118 203
CC FT DISULFID 125 194
CC FT DISULFID 130 155
CC FT DISULFID 216 245
CC FT DISULFID 226 237
CC FT DISULFID 294 328
CC FT DISULFID 374 435
CC FT DISULFID 381 408
CC FT CARBOHYD 87
CC FT CARBOHYD 135 135
CC FT CARBOHYD 154 154
CC FT CARBOHYD 186 186
CC FT CARBOHYD 195 195
CC FT CARBOHYD 232 232
CC FT CARBOHYD 239 239
CC FT CARBOHYD 260 260
CC FT CARBOHYD 274 274
CC FT CARBOHYD 293 293
CC FT CARBOHYD 299 299
CC FT CARBOHYD 329 329
CC FT CARBOHYD 336 336
CC FT CARBOHYD 352 352
CC FT CARBOHYD 382 382
CC FT CARBOHYD 388 388
CC FT CARBOHYD 392 392
CC FT CARBOHYD 398 398
CC FT CARBOHYD 401 401
CC FT CARBOHYD 438 438
CC FT CARBOHYD 454 454
CC FT CARBOHYD 602 602
CC FT CARBOHYD 607 607
CC FT CARBOHYD 616 616
CC FT CARBOHYD 628 628
CC SEQUENCE 847 AA; 96135 MW; 0A901317FD7FF2AB CRC64;

```


Query Match 87.4%; Score 2928; DB 1; Length 847;
 Best Local Similarity 86.0%; Pred. No. 8.5e-226;
 Matches 553; Conservative 28; Mismatches 40; Indels 22; Gaps 5;

QY 1 VEKMTVTVYGVVPMKEATTTTFCASDAKAYTEVHNWATHACVPTDNPQEVLENVT 60
 |||
 DB 30 VEKMTVTVYGVVPMKEATTTTFCASDAKAYTEVHNWATHACVPTDNPQEVLENVT 89

QY 61 EHFNMKNMVEOMEDIISLWDQSLKPCVKLTPLCVTLNCKDV-NATNTTDSBGTHER 119
 |||
 DB 90 EHFNMKNMVEOMEDIISLWDQSLKPCVKLTPLCVTLNCKDV-NATNTTDSBGTHER 149

QY 120 GEIKNCSFNITTSIRDEVOKEALFYKLDVYXIDNNNTSYRLISCDTSVITQACPKISFE 179
 |||
 DB 150 GEIKNCSFNITTSIRDEVOKEALFYKLDVYXIDNNNTSYRLISCDTSVITQACPKISFE 209

QY 180 PIPHYCAPAGFAILKCKDKTNGKPCKNVSTXCTGIRPVSTOLLNGLAEBEVY 239
 |||
 DB 210 PIPHYCAPAGFAILKCKDKTNGKPCKNVSTXCTGIRPVSTOLLNGLAEBEVY 269

QY 240 IRSDFNTNAKTIIVQLKESVEINCTRPNNNG-----AGDIRQACHN 281
 |||
 DB 270 IRSDFNTNAKTIIVQLKESVEINCTRPNNNG-----AGDIRQACHN 329

QY 282 ISRAKNDTLKQIVIKREOFENKTIIVFNHSSGDPBEIVMHSFNGSGEFFYCNSTOLFMS 341
 |||
 DB 330 ISRAKNDTLKQIVIKREOFENKTIIVFNHSSGDPBEIVMHSFNGSGEFFYCNSTOLFMS 389

QY 342 TWNNNTSGSNNTGNTTLPCKRIKQIINNMOEYKAMTAPPIRGQIRCSNNTGLLTRD 401
 |||
 DB 390 TWNNNTSGSNNTGNTTLPCKRIKQIINNMOEYKAMTAPPIRGQIRCSNNTGLLTRD 447

QY 402 GGIN-ENGTEIFRPGGDMRDWSELYKYKVKLEPGVATPKRRVQSEKAVGIG 460
 |||
 DB 448 GGIN-ENGTEIFRPGGDMRDWSELYKYKVKLEPGVATPKRRVQSEKAVGIG 507

QY 461 AVFLGFLGAGSTMGASNTLTVQARLLSGIVQOQNNLLRAIEAQOQLQITWNGIKQL 520
 |||
 DB 508 AVFLGFLGAGSTMGASNTLTVQARLLSGIVQOQNNLLRAIEAQOQLQITWNGIKQL 567

QY 551 QARVLAVERRYLQDQQLIGWCSGKLICTTAVPNAASNSKSLDRIWNNMTMEVEREID 580
 |||
 DB 568 QARVLAVERRYLQDQQLIGWCSGKLICTTAVPNAASNSKSLDRIWNNMTMEVEREID 627

QY 581 NYSRTYTLIESONQOEKNEQLELDKMASLWMPDITTKLWY 625
 |||
 DB 628 NYSRTYTLIESONQOEKNEQLELDKMASLWMPDITTKLWY 672

RESULT 3
 ENV_HVISC STANDARD; PRT; 856 AA.
 AC P05878;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Envelope polypeptide GP160 precursor (contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)).
 GN ENV.
 OS Human immunodeficiency virus type 1 (SC isolate) (HIV-1).
 OC Virus; Retrovirus; Retroviridae; Lentivirus.
 OC NCBI_TaxID=11702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88219542; PubMed=3369091;
 RA Gurgo C., Guo H.-G., Franchini G., Aldovini A., Collalti B.,
 RA Farrell K., Wong-Staal F., Gallo R.C., Reitz M.S. Jr.;
 RT "Envelope sequences of two new United States HIV-1 isolates."
 RL Virology 164:531-536(1988)
 CC -1- MISCELLANEOUS: THE SC ISOLATE WAS TAKEN FROM AN ARC PATIENT IN
 CC 1984 IN SOUTHERN CALIFORNIA.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL, M17450; NOT_ANNOTATED_CDS.
 DR HIV, M17450; ENV58C.
 DR InterPro: IPR000328; Env GP41.
 DR InterPro: IPR000777; GP120.
 DR Pfam: PF00516; GP120; 1.
 DR Pfam: PF00517; GP41; 1.
 KW Aids; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
 KW Signal.
 FT SIGNAL 1 29
 FT CHAIN 30 510
 FT CHAIN 511 856
 FT SITE 760 760
 FT DISULFID 53 73
 FT DISULFID 118 206
 FT DISULFID 125 197
 FT DISULFID 130 160
 FT DISULFID 219 247
 FT DISULFID 228 239
 FT DISULFID 296 330
 FT DISULFID 376 439
 FT DISULFID 383 412
 FT CARBOHYD 87 87
 FT CARBOHYD 129 129
 FT CARBOHYD 135 135
 FT CARBOHYD 140 140
 FT CARBOHYD 143 143
 FT CARBOHYD 159 159
 FT CARBOHYD 163 163
 FT CARBOHYD 189 189
 FT CARBOHYD 198 198
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 302 302
 FT CARBOHYD 331 331
 FT CARBOHYD 338 338
 FT CARBOHYD 354 354
 FT CARBOHYD 360 360
 FT CARBOHYD 384 384
 FT CARBOHYD 394 394
 FT CARBOHYD 400 400
 FT CARBOHYD 405 405
 FT CARBOHYD 442 442
 FT CARBOHYD 457 457
 FT CARBOHYD 611 611
 FT CARBOHYD 616 616
 FT CARBOHYD 625 625
 FT CARBOHYD 637 637
 FT CARBOHYD 674 674
 FT CARBOHYD 816 816
 SQ SEQUENCE 856 AA; 97055 MW; DAF4DA600EBA7A08 CRC64;

Query Match 86.6%; Score 2903.5; DB 1; Length 856;
 Best Local Similarity 84.4%; Pred. No. 7.8e-224;
 Matches 553; Conservative 31; Mismatches 36; Indels 35; Gaps 8;

QY 2 EKLMTVTVYGVVPMKEATTTTFCASDAKAYTEVHNWATHACVPTDNPQEVLENVT 61
 |||
 DB 31 EKLMTVTVYGVVPMKEATTTTFCASDAKAYTEVHNWATHACVPTDNPQEVLENVT 90

QY 62 EHFNMKNMVEOMEDIISLWDQSLKPCVKLTPLCVTLNCKDV-NATNTTDSBGTHER 115
 |||

```

Db      91 NFNMMKNNMVEQWHEDIISLMDQSLKPCVLTPLCYTLNCLNLTNDSTNATNTSSNRG 150
Qy      116 TMRGELKNCSPNTTTSIRDEVQKVALFYKLDVYXIDNNNTSRLISCTSVTQACPK 175
Db      151 KMGEGENTCSFNTTTSIRSVQKVALFYKLDVYDID--NTSTYLLNCLNNTSVTQACPK 208
Qy      176 ISPEPIHYCAPAGFAILKCNKDTFNGKXPCKNVSTXQCTGIRPVSTQLLNGSLAE 235
Db      209 VSEFPIPIHYCA-RWFAIILNCKNKGCTGCTVSTVQCTGIRPVSTQLLNGSLAE 267
Qy      236 EEVVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNGA-----GDIRQ 277
Db      268 EEVVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNTTSIHIGRAFYATGDIIGDIRQ 327
Qy      278 AHNCISAKKNDTKQIVIKLRQFENKTIIVFNHSSGDDPEIWNHSCGGEFFYCNSTQ 337
Db      328 AHNCISAKKNNNTKQIVIKLRQFENKTIIVFNHSSGDDPEIWNHSCGGEFFYCNSTQ 387
Qy      338 LFNSTNNNTGSSNTEGN--TTLPCRIKQIINNMOEGKAMYAPRIGQIRCSNITGL 396
Db      388 LFSSTW-NGTBSGNTGNDITTLPCRIKQIINNMOEGKAMYAPRIGQIRCSNITGL 446
Qy      397 LITRDGGINENG---TEIFRGGSDMRDNRSELYKXVYKIEPLGVAFTKCRVYQ 451
Db      447 LITRDGNSKXGSKNENTEIFRPGGDMRDNRSELYKXVYKIEPLGVAFTKCRVYQ 506
Qy      452 REKRAVG-IGAVFPGFGAGSTGASMTLTVOARLLSGIYOOONNLRAIEAOQRL 510
Db      507 REKRAVGTIGAMFGFGAGSTGASMTLTVOARLLSGIYOOONNLRAIEAOQRL 566
Qy      511 QLTVMGKQIOLARVLAVERYLGDQQLIGWCSGKLTCTTAVPNNASMSNKSNDRIWNNM 570
Db      567 QLTVMGKQIOLARVLAVERYLGDQQLIGWCSGKLTCTTAVPNNASMSNKSNDRIWNNM 626
Qy      571 TMMWESEIDNNTSEITLLEESONQOKBOELLEDKASLNNWDTIKMLY 625
Db      627 TMMWESEIDNNTSEITLLEESONQOKBOELLEDKASLNNWDTIKMLY 681

RESULT 4
ENV_HV1BR STANDARD; PRT; 861 AA.
ID ENV_HV1BR STANDARD; PRT; 861 AA.
AC P03377;
DT 21-JUN-1986 (Rel. 01, Created)
DT 21-JUN-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Envelope polypeptide GP160 precursor (Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)).
ENV.
OS Human immunodeficiency virus type 1 (BRU isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11686;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=8509333; PubMed=2981635;
RA Main-Hobson S., Sonigo P., Danos O., Cole S., Allion M.;
RT "Nucleotide sequence of the AIDS virus, LAV.";
RL Cell 40:9-17(1985).

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (See http://www.1db-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
DR EMBL: K02013; AAB59751.1; -
DR EMBL: A04321; CA00352.1; -
DR PIR: A03975; VCLJLV.
DR PDB: 1ERF; 20-FEB-02.
DR HIV: K02013; ENVSBRU.
DR InterPro: IPR000328; Env_GP41.

```

```

DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal;
KW 3D-structure.
FT SIGNAL 1 30
FT CHAIN 31 516
FT CHAIN 517 861
FT DISULFID 54 74
FT DISULFID 119 210
FT DISULFID 126 201
FT DISULFID 131 162
FT DISULFID 223 252
FT DISULFID 233 244
FT DISULFID 301 336
FT DISULFID 383 450
FT DISULFID 390 423
FT CARBOHYD 88 88
FT CARBOHYD 136 136
FT CARBOHYD 141 141
FT CARBOHYD 146 146
FT CARBOHYD 161 161
FT CARBOHYD 165 165
FT CARBOHYD 191 191
FT CARBOHYD 202 202
FT CARBOHYD 235 235
FT CARBOHYD 239 239
FT CARBOHYD 246 246
FT CARBOHYD 267 267
FT CARBOHYD 281 281
FT CARBOHYD 294 294
FT CARBOHYD 300 300
FT CARBOHYD 306 306
FT CARBOHYD 337 337
FT CARBOHYD 344 344
FT CARBOHYD 361 361
FT CARBOHYD 391 391
FT CARBOHYD 397 397
FT CARBOHYD 402 402
FT CARBOHYD 411 411
FT CARBOHYD 453 453
FT CARBOHYD 468 468
FT CARBOHYD 616 616
FT CARBOHYD 621 621
FT CARBOHYD 630 630
FT CARBOHYD 642 642
FT CARBOHYD 679 679
FT CARBOHYD 755 755
FT CARBOHYD 821 821
SQ SEQUENCE 861 AA; 97487 MW; 04DE2B4D4B4F6D3A CRC64;

Query Match 86.6%; Score 2901.5; DB 1; Length 861;
Best Local Similarity 84.3%; Pred. No. 1,le-223;
Matches 552; Conservative 31; Mismatches 41; Indels 31; Gaps 7;

Qy      2 EKLWVTVYGVAPWKEATTTLPFCASDAKAYDEVHANVMAHACVPTDPNQEVLNVTE 61
Db      32 EKLWVTVYGVAPWKEATTTLPFCASDAKAYDEVHANVMAHACVPTDPNQEVLNVTE 91

Qy      62 HFNMMKNNMVEQWHEDIISLMDQSLKPCVLTPLCYTLNCLNLTNDSTNATNTSSNRG 115
Db      92 NFNMMKNNMVEQWHEDIISLMDQSLKPCVLTPLCYTLNCLNLTNDSTNATNTSSNRG 151

Qy      116 TMRGELKNCSPNTTTSIRDEVQKVALFYKLDVYXIDNNNTSRLISCTSVTQACPK 174
Db      152 KMGEGENTCSFNTTTSIRSVQKVALFYKLDVYDID--NTSTYLLNCLNNTSVTQACPK 211

Qy      175 ISPEPIHYCAPAGFAILKCNKDTFNGKXPCKNVSTXQCTGIRPVSTQLLNGSLAE 234
Db      212 VSEFPIPIHYCA-RWFAIILNCKNKGCTGCTVSTVQCTGIRPVSTQLLNGSLAE 271

Qy      235 EEVVIRSDNFTNNAKTIIVOLKESVEINCTRPNN-----GAGDI 275

```

```

Db 272 EEEVIRSANFTDNKTTIIVOLNQSVETNCTRPNNTRKRSIRIORGPGRAFTTICKIGNM 331
QY 276 ROAHGNISRAKWNATLKQIVILAREOF-ENKTIIVFNHSGSGDEPEIYMSFNGCGEFPYCN 334
Db 332 ROAHGNISRAKWNATLKQIVILAREOF-ENKTIIVFNHSGSGDEPEIYMSFNGCGEFPYCN 391
QY 335 STOLFNSSTWNN--NTEGSSNTEG-NTITLPCRIOIIMMOEYKAMYPPIRQOIRCS 390
Db 392 STOLFNSSTWNNSTEGSSNTEGSDTITLPCRIOIIMMOEYKAMYPPIRQOIRCS 451
QY 391 SNITGILLTROGGINENGEIETFRPGGDMRDWSELYKYKYKIEPLGVAATCKRRV 450
Db 452 SNITGILLTROGGINENGEIETFRPGGDMRDWSELYKYKYKIEPLGVAATCKRRV 511
QY 451 OREKRAVGIGAVFLGAGSTMGASWTLFTVQRLILSGIVQOONNLRAIEAQOHL 510
Db 512 OREKRAVGIGAVFLGAGSTMGASWTLFTVQRLILSGIVQOONNLRAIEAQOHL 571
QY 511 QLTWGIKOLQARVLAVERYLGDQQLGIMGCSGKLICTTAVPMNASNSKSLDRIMNM 570
Db 572 QLTWGIKOLQARVLAVERYLGDQQLGIMGCSGKLICTTAVPMNASNSKSLDRIMNM 631
QY 571 TMMWEERIDNTSITITLIESONQOEKNEBELLDKMASLWMPDITTKLWY 625
Db 632 TMMWEERIDNTSITITLIESONQOEKNEBELLDKMASLWMPDITTKLWY 686

```

RESULT 5

ENV_HV1B1 STANDARD; PRT; 856 AA.

```

ID ENV_HV1B1 STANDARD; PRT; 856 AA.
AC P03375;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
  glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (BH10 isolate) (HIV-1).
OC Viruses; Retroviridae; Retroviridae; Lentiviridae.
OX NCBI_TaxID=11678;
RA MEDLINE=85111123; PubMed=2578615;
RA Raftery L., Haeveline W., Patarca R., Liyak K.J., Starcich B.R.,
  Joseph S.F., Doran B.R., Rafalski J.A., Whitehorn B.A.,
  Baumbach K., Ivanoff L., Peteway S.R., Pearson M.L.,
  Lautenberger J.A., Papas T.S., Grayed J., Chang N.T., Gallo R.C.,
  Wong-Staal F.;
  "Complete nucleotide sequence of the AIDS virus, HTLV-III.";
  Nature 313:277-284 (1985).

```

```

RN [2]
RP DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=90285159; PubMed=2355006;
RA Leonard C.K., Spelman M.W., Riddle L., Harris R.J., Thomas J.N.,
  Gregory T.J.;
  "Assignment of intrachain disulfide bonds and characterization of
  potential glycosylation sites of the type 1 recombinant human
  immunodeficiency virus envelope glycoprotein (gp120) expressed in
  Chinese hamster ovary cells.";
  J. Biol. Chem. 265:10373-10382 (1990).

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).

```

```

CC EMBL: M15654; AAA44205.1;
DR PIR: A03973; VCLTH3.
DR HIV: M15654; ENVSBH102.
DR InterPro: IPR000328; Env_GP41.

```

```

DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KW AIDS; Coac protein; Polyprotein; Glycoprotein; Transmembrane;
  Signal.
FT SIGNAL 1 30
FT CHAIN 31 511
FT CHAIN 512 856
FT DISULFID 54 74
FT DISULFID 119 205
FT DISULFID 126 196
FT DISULFID 131 157
FT DISULFID 218 247
FT DISULFID 228 239
FT DISULFID 296 331
FT DISULFID 378 445
FT DISULFID 385 418
FT CARBOHYD 88
FT CARBOHYD 136 136
FT CARBOHYD 141 141
FT CARBOHYD 156 156
FT CARBOHYD 160 160
FT CARBOHYD 186 186
FT CARBOHYD 197 197
FT CARBOHYD 230 230
FT CARBOHYD 234 234
FT CARBOHYD 241 241
FT CARBOHYD 262 262
FT CARBOHYD 276 276
FT CARBOHYD 289 289
FT CARBOHYD 301 301
FT CARBOHYD 301 301
FT CARBOHYD 332 332
FT CARBOHYD 339 339
FT CARBOHYD 356 356
FT CARBOHYD 386 386
FT CARBOHYD 392 392
FT CARBOHYD 397 397
FT CARBOHYD 406 406
FT CARBOHYD 448 448
FT CARBOHYD 463 463
FT CARBOHYD 611 611
FT CARBOHYD 616 616
FT CARBOHYD 625 625
FT CARBOHYD 637 637
FT CARBOHYD 674 674
FT CARBOHYD 750 750
FT CARBOHYD 816 816
SQ SEQUENCE 856 AA; 97224 MW; 0BFFB1A18931BB27 CRC64;

```

Query Match 86.2%; Score 2887; DB 1; Length 856;
 Bees Local Similarity 84.0%; Pred. No. 1.6e-222;
 Matches 546; Conservative 34; Mismatches 44; Indels 26; Gaps 5;

```

QY 2 EKLWTVYVYGVVWKEATTTTLCASDAKAYTEVYVWVATACVPTDPNPOEVLENTTE 61
Db 32 EKLWTVYVYGVVWKEATTTTLCASDAKAYTEVYVWVATACVPTDPNPOEVLENTTE 91
QY 62 HFNWKNMNVQWQWEDIIISLDQSLKPCVKLTPLCVTLNCKQVNAVNTTNDSEG--TWER 119
Db 92 HFNWKNMNVQWQWEDIIISLDQSLKPCVKLTPLCVTLNCKQVNAVNTTNDSEG--TWER 151
QY 120 GEIKKCSNITTSIRDBYQKRYALFYKLDVXKIDNNNTSYLLISGDTSVITQACPKISFE 179
Db 152 GEIKKCSNITTSIRDBYQKRYALFYKLDVXKIDNNNTSYLLISGDTSVITQACPKISFE 211
QY 180 PIPHYCAPAGFALIKCDKTFNGCKPCKNVSTXQCTHGIRPVVSTQLLNGSLAEERVV 239
Db 212 PIPHYCAPAGFALIKCDKTFNGCKPCKNVSTXQCTHGIRPVVSTQLLNGSLAEERVV 271
QY 240 IRSDFNNAKTTIIVOLKESVEINCTRPNNN-----GAGDIRQAH 280
Db 272 IRSANFTDNKTTIIVOLNQSVETNCTRPNNNTRKRSIRIORGPGRAFTTICKIGNMRAHC 331

```

QY 281 NISRAKNDTLKOIVIKLREOF-ENKTIIVENHSSGDPPEIYVHSGGEPFYCNSTOLF 339
 DB 332 NISRAKNNLTKOJDSKLRDFGNKKTIIIFKQSSGDPPEIYVHSGGEPFYCNSTOLF 391
 QY 340 NSTWNN--NTEGSNTEG-NTITLPCRIOIIMMOEYGVKAWAPPRIQIRCSSNITG 395
 DB 392 NSTWNNSTWSTKSGSNTEGSDITLPCRIOIIMMOEYGVKAWAPPRIQIRCSSNITG 451
 QY 396 LLTRDGIENGEIIRPGGDMRDMWRSSELYKYVVKIIEPLGVAATKCRARVOREKR 455
 DB 452 LLTRDGSNNSEIIRPGGDMRDMWRSSELYKYVVKIIEPLGVAATKCRARVOREKR 511
 QY 456 AVIGAVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOONLLRAIEAOQRLQTLTW 515
 DB 512 AVIGAVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOONLLRAIEAOQRLQTLTW 571
 QY 516 GIKOLQARLAVERYLDQQLIGWCSGKLICTTAVPWNASWNSKSLDRIMNMWMEW 575
 DB 572 GIKOLQARLAVERYLDQQLIGWCSGKLICTTAVPWNASWNSKSLDRIMNMWMEW 631
 QY 576 EREIDNTSEIYTIIESONOEKOELELDKMAIIMWFDITKMLY 625
 DB 632 EREIDNTSEIYTIIESONOEKOELELDKMAIIMWFDITKMLY 681
 RESULT 6
 ENV_HV1J3 STANDARD; PRT; 867 AA.
 AC P12489;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Envelope glycoprotein GP160 precursor [contains: Exterior membrane
 glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
 GN ENV.
 OS Human immunodeficiency virus type 1 (HIV isolate) (HIV-1).
 OC Viruses; Retroviruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11694;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=69352108; PubMed=2669897;
 RA Komiyama N., Hattori N., Inoue J., Sakuma S., Kurimura T., Yoshida M.,
 "Nucleotide sequences of gag and env genes of a Japanese isolate of
 HIV-1 and their expression in bacteria."
 RT AIDS Res. Hum. Retroviruses 5:411-419 (1989).
 RL AIDS Res. Hum. Retroviruses 5:411-419 (1989).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M21138; AAB03526.1; -
 DR HIV: M21138; ENV.SJ3.
 DR InterPro: IPR000328; Env_GP41.
 DR InterPro: IPR000777; GP120.
 DR Pfam: PF00516; GP120; 1.
 DR Pfam: PF00517; GP41; 1.
 KM AIDS: Coat protein; Polypeptide; Glycoprotein; Transmembrane;
 KW signal.
 FT SIGNAL. 1 30
 FT CHAIN 31 516 EXTERIOR MEMBRANE GLYCOPROTEIN.
 FT CHAIN 517 867 TRANSMEMBRANE GLYCOPROTEIN.
 FT DISULFID 53 73 BY SIMILARITY.
 FT DISULFID 118 217 BY SIMILARITY.
 FT DISULFID 125 208 BY SIMILARITY.
 FT DISULFID 130 160 BY SIMILARITY.
 FT DISULFID 230 259 BY SIMILARITY.
 FT DISULFID 240 251 BY SIMILARITY.
 FT DISULFID 308 342 BY SIMILARITY.

FT DISULFID 388 457 BY SIMILARITY.
 FT DISULFID 395 430 BY SIMILARITY.
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 135 135 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 140 140 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 143 143 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 159 159 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 163 163 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 188 188 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 189 189 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 199 199 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 209 209 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 246 246 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 253 253 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 274 274 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 288 288 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 307 307 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 350 350 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 366 366 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 372 372 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 396 396 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 408 408 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 412 412 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 418 418 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 423 423 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 460 460 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 475 475 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 622 622 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 627 627 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 636 636 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 648 648 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 867 AA; 98399 MW; 5F23101468B8680 CRC64;
 Query Match 86.1%; Score 2884; DB 1; Length 867;
 Best Local Similarity 83.6%; Pred. No. 2,9e-222;
 Matches 554; Conservat 25; Mismatches 44; Indels 40; Gaps 6;
 QY 2 EKLMTVTVYGVVPWKATTTLPFCASDAKAYDTEVHNVMAHACVPDPNPOEVLNVT 61
 DB 31 EQLMTVTVYGVVPWKATTTLPFCASDAKAYDTEVHNVMAHACVPDPNPOEVLNVT 90
 QY 62 HFNWKKNNWVQOEEDIIISLMDQSLKPCVLTPLCVTLNCKD-----VNAITNTDSEG 115
 DB 91 KFNWKKNNWVQOEEDIIISLMDQSLKPCVLTPLCVTLNCKD-----VNAITNTDSEG 150
 QY 116 TMRGEIKGCSFNITTSIRDEVOKEYALFYKLDVYXIDN-----NNTSRLISCDT 166
 DB 151 KMRGEIKGCSFNITTSIRDEVOKEYALFYKLDVYXIDN-----NNTSRLISCDT 210
 QY 167 SVITQACPRTSPFPIHYCAPGAFATLKGNKDPNFKXPCKNVSTQCTHGRIPVSTQ 226
 DB 211 SVITQACPRTSPFPIHYCAPGAFATLKGNKDPNFKXPCKNVSTQCTHGRIPVSTQ 270
 QY 227 LLNGLSLAESEVYIRSDNFTNNAKTIIVQLKESVEINCTRPNNNG----- 271
 DB 271 LLNGLSLAESEVYIRSDNFTNNAKTIIVQLKESVEINCTRPNNNG----- 330
 QY 272 ---AGDIRQAHCNISRAKNDTLKOIVIKLREOFENKTIIVHNSGGDPPEIYVHSGG 328
 DB 331 KOIAGDIRQAHCNINRAKNDTLKOIVIKLREOFENKTIIVHNSGGDPPEIYVHSGG 390
 QY 329 EFPYCNSTOLFNSNWTN-----NTEGSNTEG-NTITLPCRIOIIMMOEYGVKAWAPP 382
 DB 391 EFPYCNSTOLFNSNWTN-----NTEGSNTEG-NTITLPCRIOIIMMOEYGVKAWAPP 450
 QY 383 IRGQIRCSSNITGLLTRDGIENGEIIRPGGDMRDMWRSSELYKYVVKIIEPLGVA 442
 DB 451 IREGIRCSSNITGLLTRDGIENGEIIRPGGDMRDMWRSSELYKYVVKIIEPLGVA 510
 QY 443 TKCRARVOREKRAVGIGAVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOONLLRA 502
 DB 511 TKCRARVOREKRAVGIGAVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOONLLRA 569

FT	CARBOHYD	389	389	N-LINKED (GLCNAC. .)	(POTENTIAL. .)
FT	CARBOHYD	395	395	N-LINKED (GLCNAC. .)	(POTENTIAL. .)
FT	CARBOHYD	400	400	N-LINKED (GLCNAC. .)	(POTENTIAL. .)
FT	CARBOHYD	435	435	N-LINKED (GLCNAC. .)	(POTENTIAL. .)
FT	CARBOHYD	450	450	N-LINKED (GLCNAC. .)	(POTENTIAL. .)
FT	CARBOHYD	598	598	N-LINKED (GLCNAC. .)	(POTENTIAL. .)
FT	CARBOHYD	603	603	N-LINKED (GLCNAC. .)	(POTENTIAL. .)
FT	CARBOHYD	612	612	N-LINKED (GLCNAC. .)	(POTENTIAL. .)
FT	CARBOHYD	624	624	N-LINKED (GLCNAC. .)	(POTENTIAL. .)
FT	CARBOHYD	803	803	N-LINKED (GLCNAC. .)	(POTENTIAL. .)
SQ	SEQUENCE	843 AA;	95648 MW; C69DEPD71C918B71 CRC64;		
<hr/>					
Query Match 86.0% Score 2882; DB 1; Length 843;					
Best Local Similarity 85.1%; Pred. No. 4e-22;					
Matches 548; Conservative 28; Mismatches 42; Indels 26; Gaps 7;					
<hr/>					
Qy	2	EKLAVTVYGVVPEWKEATTTLFCASDAKAYDTEVHNWATHACVPDPNPOEVLNVTE	61		
Db	31	EQLMVTYYGVVPWKETITTLFCSADAKAYDTEVHNWATHACVPDPNPOEVLNVTE	90		
Qy	62	HFNKKNNWYBQNQEDTISLMDSLKPCVKLTPLCTLVCKDV-NATNTNDSEGTMERG	120		
Db	91	NFNMKNNWYBQNQEDTISLMDSLKPCVKLTPLCTLVCKDNATNTNDSEGTMERG	150		
Qy	121	EIKKCSNITTSIRDEVOKEXALPYKLDVYXINNNSYSLSDSVITIQACPKSFEF	180		
Db	151	EIKKCSNITTSIRDKVOKEXALTFLNDVPID--NASYSLISNLSVITIQACPKSFEF	208		
Qy	181	IPIHYCAPAGFAILKCNDKTFFNGKCPKANSTXOCHTGIRPVYSTOLLNGSLAEEVI	240		
Db	209	IPIHYCAPAGFAILKCNDKDFNGGTCPTNVSTVOCHTGIRPVYSTOLLNGSLAEEVI	268		
Qy	241	RSDNFVNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRAHNCI	282		
Db	269	RSENFVNNAKTIIVOLNESVAINCTRPNNTTRKSNINGPRALYTIGEIIIDIRAHNCI	328		
Qy	283	SRAMNDTLKOIYVTKLEOF-ENKTIYFNHSSGGDEPIWHSFNGGEFPYCNSTOLENS	341		
Db	329	SKTQWENTLKOIAKLEQGANNKTIIFNPSSGGDEPIVTHSFGGGEFPYCNSTOLE-	386		
Qy	342	TANNNTSGSNSNTEGNTTTLFCRIKQIIMNMQEVGAAYAPPIRGQICRSSNITGLLTRD	401		
Db	387	TW-NDTKLANIT-GRNTTLFCRIKQIIMNMQEVGAAYAPPIRGQICRSSNITGLLTRD	444		
Qy	402	GGINENGTEIPFGGDMRDNRMSLEYKYVAKIEPLGVAPTYCKRRVRQEKRAVGIGA	461		
Db	445	GGRKTNGETEIPFGCGDMRDNRMSLEYKYVAKIEPLGVAPTYAKRRVRQEKRAVGIGA	504		
Qy	462	VFLGFGLAAGSTWAAAMTLTYVARLLSGIVOOONNLRAIEAQOMQLTWIGIKOLO	521		
Db	505	LFLGFGLAAGSTWAAASITTLTYVARLLSGIVOOONNLRAIEAQOMQLTWIGIKOLO	564		
Qy	522	ARYLAVERRYLGDOOLLAGIWGCSGGLICTYAVPMNASWSKSLDRIVNMMTMWEEREIDN	581		
Db	565	ARYLAVERRYLRDQOLLGIWGCSGLICTITTVPAINTSNTSKSLNEIWMNTMMKREERIDN	624		
Qy	582	YTSEIYTLIESQNOQEKNQOEELLEDKMASLWNWFDTIKWLNY	625		
Db	625	YTHIITYSLIESQNOQEKNQOEELLALDDKMASLWNWFDTIKWLNY	668		
<hr/>					
RESULT 9					
ENV_HVJ1BN	STANDARD;	PRT;	852 AA.		
ID_P12488;					
DT	01-OCT-1989 (Rel. 12, Created)				
DT	01-OCT-1989 (Rel. 12, Last sequence update)				
DT	15-SEP-2003 (Rel. 42, Last annotation update)				
DE	Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].				
GN	ENV.				
OS	Human immunodeficiency virus type 1 (BR isolate) (HIV-1).				
OC	Viruses; Retroviral viruses; Retroviridae; Lentivirus.				

OK NCBI_TaxID=11693;
 RN SEQUENCE FROM N.A.
 RP MEDLINE=89085613; PubMed=2789516;
 RX Anand R., Thayer R., Srinivasan A., Nayar S., Gardner M., Luciw P.,
 RA Dandekar S.;
 RT "Biological and molecular characterization of human immunodeficiency
 virus (HIV-1B) from the brain of a patient with progressive
 RT dementia";
 RL Virology 168:79-89(1989).
 CC -1- MISCELLANEOUS: THIS VIRUS IS CYTOPATHICALLY ACTIVE AND WAS
 CC HARVESTED FROM THE BRAIN TISSUE OF A NEUROLOGICAL AIDS PATIENT.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M21098; AAA44221.1; -
 DR PIR: A1667; VCLJBR.
 DR PDB: 1IM7; 23-OCT-02.
 DR HIV; M21098; ENV5BRVA.
 DR InterPro; IPR000328; Env GP41.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120.1.
 DR Pfam; PF00517; GP41.1.
 KM AIDS; Coat protein; Glycoprotein; Transmembrane; Signal;
 KW 3D-structure.
 FT SIGNAL 1 30
 FT CHAIN 31 507
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 155
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 330
 FT DISULFID 376 439
 FT DISULFID 383 412
 FT CARBOHYD 49 88
 FT CARBOHYD 135 135
 FT CARBOHYD 138 138
 FT CARBOHYD 154 154
 FT CARBOHYD 158 158
 FT CARBOHYD 197 197
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 331 331
 FT CARBOHYD 354 354
 FT CARBOHYD 360 360
 FT CARBOHYD 384 384
 FT CARBOHYD 390 390
 FT CARBOHYD 396 396
 FT CARBOHYD 400 400
 FT CARBOHYD 442 442
 FT CARBOHYD 456 456
 FT CARBOHYD 607 607
 FT CARBOHYD 612 612
 FT CARBOHYD 621 621
 FT CARBOHYD 633 633
 FT CARBOHYD 670 670
 FT CARBOHYD 812 812
 SQ SEQUENCE 852 AA; 97203 MW; 2BB866345DEC915F CRC64;

Query Match 85.9%; Score 2878; DB 1; Length 852;
 Best Local Similarity 83.8%; Pred. No. 8-5e-223;
 Matches 542; Conservative 38; Mismatches 43; Indels 24; Gaps 5;
 QY 2 EKLAVTVYVGVVWKEATTTLPFCASDAKAYDTEVHNWATNACVPTDPNPOEVLENVTE 61
 DB 32 DKLAVTVYVGVVWKEATTTLPFCASDAKAYDTEVHNWATNACVPTDPNPOEVLENVTE 91
 QY 62 HNNWKNMVMOMODIISLDOSLKPCVKLTPLCVTLNCHDFNATNATNSGKMEGGE 121
 DB 92 HNNWKNMVMOMODIISLDOSLKPCVKLTPLCVTLNCHDFNATNATNSGKMEGGE 151
 QY 122 IKNCSEFNITTSIRDEQKEVALFYLDVXIDNN--NTSYRLISCDTSVITQACPISPE 179
 DB 132 MKNCSEFNITTSIRDMQKEVALFYLDVPIPDNDKTNRYRLISCVITQACRKYTPE 211
 QY 180 PIPHYCAPAGPAILKNDKTFNGKPKKXVSTQXCTHGIRPVYSTOLLNGLAEEVY 239
 DB 212 PIPHYCAPAGPAILKNDKTFNGKPKKXVSTQXCTHGIRPVYSTOLLNGLAEEVY 271
 QY 240 IRSDNFTNNACTIIVQLKESEVEINCTRPNNNG-----AGDIRQACN 281
 DB 272 IRSENFNNVKTIIIVQLNESEVEINCTRPNNNTRKRITNGPGRVYTTGQIIGDIRAHGN 331
 QY 282 ISRAKMDTLKQIVIKLEOFENKTIYFNHSSGGDEIYVHSHFNGGEPFYCNSTOLFNS 341
 DB 332 LSRSKMNTLQIVYKLVQVQKNTIYFNHSSGGDEIYVHSHFNGGEPFYCNSTOLFNS 391
 QY 342 TWNNTEGSSNNTGNT-ITLPCRKIQIINMQEYKAYAPPIRQICSSNITGLLTYR 400
 DB 392 TWYRNTTG-NITEGNSPITLPCRKIQIINMQEYKAYAPPIRQICSSNITGLLTYR 450
 QY 401 DQGINE--NGEIRPFGGDMRDNRSELYKVKIEPLGVAAPKCRVYVQREKAVG 458
 DB 451 DQGNNTTDTDEIRFPGGDMRDNRSELYKVKIEPLGVAAPKCRVYVQREKAVG 510
 QY 459 IGAVELGFLGAAGSTMAASMTLTVQARLLASGIYQOONNLLRAEAOQRMLOLTVMGIK 518
 DB 511 LGALPLGLGAGSTMAASMTLTVQARLLASGIYQOONNLLRAEAOQRMLOLTVMGIK 570
 QY 519 QLOARVLAVERVYLGDOQLGIMGSGKLICTTAVPMNASWSNKSIDRIIMNNTMWEWERE 578
 DB 571 QLOARVLAVERVYLGDOQLGIMGSGKLICTTAVPMNASWSNKSIDRIIMNNTMWEWERE 630
 QY 579 IDNTSEIYTLIESQOQEKQDELELDKVASIAMPFITKMLY 625
 DB 631 IDNTNLIVSLIEDSQIOEKKEKLELDKVASIAMPFITKMLY 677
 RESULT 10
 ENV_HV1H2 STANDARD; PRT; 856 AA.
 AC P04578; 009779;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Envelope polypeptide precursor (Contains: Exterior membrane
 DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)).
 GN ENV.
 OS Human immunodeficiency virus type 1 (HXB2 isolate) (HIV-1).
 OC Viruses; Retroid viruses; Retroviridae; Lentivirinae.
 OC NCBI_TaxID=11706;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87299196; PubMed=3040055;
 RA Ratner L., Fisher A., Jagodzinski L.L., Mitsuuya H., Liou R.-S.,
 RA Gallo R.C., Wong-Staal P.;
 RT "Complete nucleotide sequences of functional clones of the AIDS
 RT virus";
 RL AIDS Res. Hum. Retroviruses 3:57-69(1987).
 RN [2]
 RP REVISIONS.

Seq	SEQUENCE	856 AA;	97212 MW;	6FBA16AF85107FE0 CFC64;
Query Match	85.9%;	Score 2878;	DB 1;	Length 856;
Best Local Similarity	83.7%;	Prod. No. 8.5e-22;		
Matches 544;	Conservative 35;	Mismatches 45;	Indels 26;	Gaps 5
QY	2 EKLWVTVYGVGWPKWEATTTLPFCASDAKAYTEVHNWATHACVPTDENPQEVLEANTY	61		
DB	32 EKLWVTVYGVGWPKWEATTTLPFCASDAKAYTEVHNWATHACVPTDENPQEVLEANTY	91		
QY	62 HFNWKNKNMVEQWQEDIIISLWDQSLKPCVKLTPLCVTLNCKDVNATNTNDSG--TWER	119		
DB	92 HFNWKNKNMVEQWQEDIIISLWDQSLKPCVKLTPLCVSLKCTDLKNDITNTSSSGMIMBK	151		
QY	120 GEINCSFNITTSIRDEYQKAYALFYKLDVXVINDNNMVSYLLSDTSVITQACKISFE	179		
DB	152 GEINCSFNITTSIRGRQKAYEAFYKLDIIIPIDDTTSYKLTSSNTSVITQACPKVSE	211		
QY	180 PIPHYCAPAGFALIKNDKTFNGKPCPKNVTXQCTHGIRPVSTOLLNGSLAESEV	239		
DB	212 PIPHYCAPAGFALIKNDKTFNGKPCPKNVTXQCTHGIRPVSTOLLNGSLAESEV	271		
QY	240 IRSDFNTNATIIIVOLKESVEINCTREBNN-----GAGDIRQAH	280		
DB	272 IRSDFNTNATIIIVOLKESVEINCTREBNNTRKIRIRQGRGAFVITIGKIGNRQAH	331		
QY	281 NISPAKNNDTLKQIVIKLRQEF-EKTIIVFNHSSGQDEIYMHSFNGCGEFPYCNSTOLF	339		
DB	332 NISPAKNNDTLKQIVIKLRQEF-EKTIIVFNHSSGQDEIYMHSFNGCGEFPYCNSTOLF	391		
QY	340 NSTANN--NTEGSNTEG-NTITLPCIKIINNMQVGVKAMPYPPRGQIRCSNITG	395		
DB	392 NSTFNSTWSTEGSNTEGSDITLPCIKIINNMQVGVKAMPYPPRGQIRCSNITG	451		
QY	396 LLTLTRDGINNGTEIFPRGGGDMRDMNRSELYKKYKIEPLGVAPTKRRVQREKR	455		
DB	452 LLTLTRDGINNGTEIFPRGGGDMRDMNRSELYKKYKIEPLGVAPTKRRVQREKR	511		
QY	456 AVGIGAVFLGFLGAAGSTMGASNTLYQARLLSGIVQOONNLRAIEAQRMQLQTYW	515		
DB	512 AVGIGAVFLGFLGAAGSTMGASNTLYQARLLSGIVQOONNLRAIEAQRMQLQTYW	571		
QY	516 GIKLOARVLAVERVYLGQQLLGIVGCGSKLICCAVPWNASWSKSLDIRINNNTMWEW	575		
DB	572 GIKLOARVLAVERVYLGQQLLGIVGCGSKLICCAVPWNASWSKSLDIRINNNTMWEW	631		
QY	576 EREIDNYSSEIYTLIEESQNOEQENQEBELDKVASLMMNFDTIKWLMY	625		
DB	632 DREINNTYSLHSLIEESQNOEQENQEBELDKVASLMMNFDTIKWLMY	681		
RESULT 11				
ENV_HV1H3				
ID ENV_HV1H3	STANDARD;	PRT;	856 AA.	
AC	P04624;			
DT	13-AUG-1987 (Rel. 05, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].			
GN	ENV.			
OS	Human immunodeficiency virus type 1 (HXB3 isolate) (HIV-1).			
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.			
OX	NCBI_TaxID=11707;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=8528248; PubMed=2988795;			
RA	Crowl R., Ganguly K., Gordon M., Conroy R., Schaber M., Kramer R.,			
RA	Shaw G.M., Wong-Staal F., Reddy E.P.;			
RT	"HIV-1: env gene products synthesized in E. coli are recognized by			
RL	antibodies present in the sera of AIDS patients.";			
CC	Cell 41:979-986(1985).			

DR PDB; INJO; 25-FEB-03.
 DR HIV; M17449; ENVSMN.
 DR InterPro; IPR000328; Env GP41.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 DR AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal;
 3D-structure.
 KM SIGNAL 1 29
 FT CHAIN 30 513 EXTERIOR MEMBRANE GLYCOPROTEIN.
 FT CHAIN 514 856 TRANSMEMBRANE GLYCOPROTEIN.
 FT DISULFID 53 73 BY SIMILARITY.
 FT DISULFID 118 210 BY SIMILARITY.
 FT DISULFID 125 201 BY SIMILARITY.
 FT DISULFID 130 162 BY SIMILARITY.
 FT DISULFID 223 252 BY SIMILARITY.
 FT DISULFID 233 244 BY SIMILARITY.
 FT DISULFID 301 335 BY SIMILARITY.
 FT DISULFID 381 445 BY SIMILARITY.
 FT DISULFID 388 418 BY SIMILARITY.
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 129 129 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 140 140 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 146 146 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 161 161 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 165 165 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 191 191 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 202 202 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 267 267 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 281 281 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 343 343 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 359 359 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 365 365 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 395 395 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 405 405 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 413 413 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 465 465 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 612 612 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 617 617 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 626 626 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 638 638 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 856 AA; 97140 MW; D197B0809940BE732 CRC64;
 Query Match 85.6%; Score 2870; DB 1; Length 856;
 Best Local Similarity 83.4%; Pred. No. 3.7e-22;
 Matches 548; Conservative 29; Mismatches 42; Indels 38; Gaps 7;

QY 234 AAEVVRISDNFTNNAKTIIVOLKESVEINCTBN-----NNGADI 275
 DB 271 AAEVVRISDNFTNNAKTIIVHLMNESVQINCTBNPKRKRIHIGPRAFYTTNIGTI 330
 QY 276 ROAHCHISRAKNDTKQIYIKREOPENKTIYPNHSGGDPPIVHNSFNCGEFFCNS 335
 DB 331 ROAHCHISRAKNDTKQIYIKREOPENKTIYPNHSGGDPPIVHNSFNCGEFFCNS 390
 QY 336 TOLFNSITWN-----NNEGSNTEGNTITLPCRIKOIINMOEYGVAMPPIRGORCS 390
 DB 391 SPLFNSTWNGNNTWNTGNS---NNITLQCKIKQIINMOEYGVAMPPIRGORCS 446
 QY 391 SNITGLLTREDCG--INENGTEIFRPGGDMRDMRSELYKYVVKIEPVGVAFTCKRR 448
 DB 447 SNITGLLTREDCGMDTNDTEIFRPGGDMRDMRSELYKYVVKIEPVGVAFTCKRR 506
 QY 449 VVOREKRAVIGAVFLGFLGAAGSTWKAASMTLTVQARLLSGIVQOQNNLRAIEQOR 508
 DB 507 VVOREKRA-AIGALFLGFLGAAGSTWKAASMTLTVQARLLSGIVQOQNNLRAIEQOR 565
 QY 509 MLOLTWGIKOLARVLAVERVIGDQOLGIMGSGKLICTAVPNNASMSKSLDIRYN 568
 DB 566 MLOLTWGIKOLARVLAVERVIGDQOLGIMGSGKLICTAVPNNASMSKSLDIRYN 625
 QY 569 NMTWMEEREIDNVTSEIYTLIEBSONQOEKNEDELLEDKWASLWNPDIITKWLWY 625
 DB 626 NMTWMEEREIDNVTSEIYTLIEBSONQOEKNEDELLEDKWASLWNPDIITKWLWY 682

RESULT 13
 ENV_HVILW STANDARD; PRT; 856 AA.
 AC Q70626;
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
 GN ENV.
 OS Human immunodeficiency virus type 1 (TM12.3 isolate) (HIV-1).
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.
 OC NCBI_TaxID=82834;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95127297; PubMed=7826699;
 RA Reitz M.S. Jr., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.M., Shaw G.M., Kong L.T., Weiss S.H., Waters D., Gallo R.C., Blattner W.;
 RT "Viral variability and serum antibody response in a laboratory worker infected with HIV type 1 (HIV type 11B).";
 RL AIDS Res. Hum. Retroviruses 10:1143-1155(1994).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ac.uk/announce/> or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL; J12053; AA76690.1; -
 DR PDB; 1IF3; 02-MAY-01.
 DR GLCOSULEDB; Q70626; -
 DR InterPro; IPR000328; Env GP41.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal;
 3D-structure.
 FT SIGNAL 1 30
 FT CHAIN 31 511 EXTERIOR MEMBRANE GLYCOPROTEIN.
 FT CHAIN 512 856 TRANSMEMBRANE GLYCOPROTEIN.
 FT DISULFID 54 74 BY SIMILARITY.
 FT DISULFID 119 205 BY SIMILARITY.

FT	DISULFID	126	196	BY SIMILARITY.
FT	DISULFID	131	157	BY SIMILARITY.
FT	DISULFID	218	247	BY SIMILARITY.
FT	DISULFID	228	239	BY SIMILARITY.
FT	DISULFID	296	331	BY SIMILARITY.
FT	DISULFID	378	445	BY SIMILARITY.
FT	DISULFID	385	418	BY SIMILARITY.
FT	CARBOHYD	88		N-LINKED (GLCNAC. -.) (POTENTIAL).
FT	CARBOHYD	136	136	N-LINKED (GLCNAC. -.) (POTENTIAL).
FT	CARBOHYD	141	141	N-LINKED (GLCNAC. -.) (POTENTIAL).
FT	CARBOHYD	156	156	N-LINKED (GLCNAC. -.) (POTENTIAL).
FT	CARBOHYD	160	160	N-LINKED (GLCNAC. -.) (POTENTIAL).
FT	CARBOHYD	186	186	N-LINKED (GLCNAC. -.) (POTENTIAL).
FT	CARBOHYD	197	197	N-LINKED (GLCNAC. -.) (POTENTIAL).
FT	CARBOHYD	230	230	N-LINKED (GLCNAC. -.) (POTENTIAL).
FT	CARBOHYD	234	234	N-LINKED (GLCNAC. -.) (POTENTIAL).
FT	CARBOHYD	241	241	N-LINKED (GLCNAC. -.) (POTENTIAL).
FT	CARBOHYD	262	262	N-LINKED (GLCNAC. -.) (POTENTIAL).
FT	CARBOHYD	276	276	N-LINKED (GLCNAC. -.) (POTENTIAL).
FT	CARBOHYD	289	289	N-LINKED (GLCNAC. -.) (POTENTIAL).
FT	CARBOHYD	295	295	N-LINKED (GLCNAC. -.) (POTENTIAL).
FT	CARBOHYD	301	301	N-LINKED (GLCNAC. -.) (POTENTIAL).
FT	CARBOHYD	332	332	N-LINKED (GLCNAC. -.) (POTENTIAL).
FT	CARBOHYD	339	339	N-LINKED (GLCNAC. -.) (POTENTIAL).
FT	CARBOHYD	356	356	N-LINKED (GLCNAC. -.) (POTENTIAL).
FT	CARBOHYD	386	386	N-LINKED (GLCNAC. -.) (POTENTIAL).
FT	CARBOHYD	392	392	N-LINKED (GLCNAC. -.) (POTENTIAL).
FT	CARBOHYD	406	406	N-LINKED (GLCNAC. -.) (POTENTIAL).
FT	CARBOHYD	448	448	N-LINKED (GLCNAC. -.) (POTENTIAL).
FT	CARBOHYD	463	463	N-LINKED (GLCNAC. -.) (POTENTIAL).
FT	CARBOHYD	611	611	N-LINKED (GLCNAC. -.) (POTENTIAL).
FT	CARBOHYD	616	616	N-LINKED (GLCNAC. -.) (POTENTIAL).
FT	CARBOHYD	624	624	N-LINKED (GLCNAC. -.) (POTENTIAL).
FT	CARBOHYD	637	637	N-LINKED (GLCNAC. -.) (POTENTIAL).
FT	CARBOHYD	674	674	N-LINKED (GLCNAC. -.) (POTENTIAL).
FT	CARBOHYD	750	750	N-LINKED (GLCNAC. -.) (POTENTIAL).
FT	CARBOHYD	816	816	N-LINKED (GLCNAC. -.) (POTENTIAL).
50	SEQUENCE	856 AA;	96938 MW;	0C241332CF7E6687 CR664;

Query Match	Similarity	85.5%	Pred	2864	DB 1	Length	856
Best Local	Similarity	85.5%	Pred	No. 1.1e-220			
Matches	543	Conservative	36	Mismatches	45	Indels	26
							Gaps
							6
QY	2	EKLWTVYGVVPMKEATTTTLCASDAKAYDTEVHNVATHACVPTDPNPQVEVLENTYE	61				
Db	32	EKLWTVYGVVPMKEATTTTLCASDAKAYDTEVHNVATHACVPTDPNPQVEVLYNTE	91				
	62	HNFMKKNMVGOMODDIISLMDQSLKPECVKLTPLCVTLNKGCV-NATNTNDSEG-THER	119				
	92	NFMNKKMDVMQMHEDIISLMDQSLKPCVKLTPLCVSLKCTDKNDNTNNSGGIMEX	151				
QY	120	GEINKCSFNITTSIRDEVOKEALPEYKLDVVYIDNNNNSYRLISGDTSYITACPKISPE	179				
Db	152	GEINKCSFNITTSIRGQYKEXAFEPYKDDIIPIDNDTTSYTLTSCNTSYITACRVSPE	211				
QY	180	PIPIHYCAPAGFALIKNDKDTFNGKPCCKNSTXOCTGIRPVSTOLLNGLAAEEVY	239				
Db	212	PIPIHYCAPAGFALIKCNKNTFNGGCPCTNVSTVOCTGIRKPVSTOLLNGLAAEEVY	271				
QY	240	IRSDNFTNNAATIIYQLKESVEINSTRPNNN-----GAGDILROAHC	280				
Db	272	IRSANLIDNVKTIIVQLNQSVEINSTRPNNTNRKRIIRORGGRFTVYIGKIGNRROAHC	331				
QY	281	NISPAKNDTLKQIVIKLREOF-ENKTIIVFNHSSGADPEIVMHSFNCGEFFYCNSTOLF	339				
Db	332	NISPAKNNNTIKQIASKLREOYGNNKTIIFKQSSGDELEIVTHSFNCGEFFYCNSTOLF	391				
QY	340	NSTYNN--NTEGNSNTEG-NITTLPCRIKQIINNMOEGKMYAPPIRGQIRCSSNNTG	395				
Db	392	NSTYFNSTWSTEGSNTEGSDTITLPCRIKQIINNMOEGKMYAPPIISGQIRCSSNNTG	451				
QY	396	LLLRDGINENGTEIIFRPGGDMDMNRSELKYKVKVIEPLGVAPTKCKRRVQREKR	455				

Dd		452	LLLRDGGANNNGSIFRPGGGDMWDRSSLYKTKVXIPDLGVAPTKARRVVQREKR	511
Oy		456	AVGICGAFLGLTGAAGSTMGAASTMLTIVQARLLLSGIYQQOQNNLRAIEAQQRMLOLTVM	515
Dd		512	AVGIGALFGLTGLGAAGSTMGAASTMTLTYQANQLISGIYQQOQNNLRAIEAQHLLQLTVM	571
Oy		516	GIKOLQARVLTAVERYLGDQOLLGWIGSGSKLICCTPAVPNNAWSNKSIDRIINNNTMEWM	575
Dd		572	GIKOLOARIILTAVERYYLKXQQOLLGWIGSGSKLICCTPAVPNNAWSNKSLEOIIMHTTMEM	631
Oy		576	EKEIDNTSELYTLIEESONOENKEOEELFDKPAISLMNPDTITKMAY	625
Dd		632	DREINNTSLTHLSIESONQEOROEELLFDKPASIMNPNITNMPLY	681
 RESULT 14 ENV_HVI1B8				
ID	_ENV_HVI1B8	STANDARD,	PRT;	851 AA.
AC	P04582;			
DT	13-AUG-1987	(Rel. 05, Created)		
DT	13-AUG-1987	(Rel. 05, Last sequence update)		
DT	15-SEP-2003	(Rel. 42, Last annotation update)		
DE	Envelope polyprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].			
GN	ENV.			
OS	Human immunodeficiency virus type 1 (BH6 isolate) (HIV-1).			
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.			
OX	NCBI_TaxID=11684;			
RX	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=85111123; PubMed=2578615;			
RA	Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R., Joseph S.F., Doran E.R., Rafalski J.A., Whitehorn E.A., Baumeister K., Ivanoff L., Pettenay S.R., Jr., Pearson M.L., Lautenberger J.A., Papas T.S., Ghrayed Y., Chang N.T., Gallo R.C., Wong-Staal F.;			
RL	"Complete nucleotide sequence of the AIDS virus, HTLV-III.";			
RT	Nature 313:277-284(1985).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; KO2011; AAA44661.1; .			
DR	PDB; 1AIK; 16-JUN-97.			
DR	PDB; 1DDH; 13-JAN-99.			
DR	PDB; 1HHG; 31-OCT-93.			
DR	PDB; 1QO3; 02-JAN-00.			
DR	PDB; 1SZT; 24-DEC-97.			
DR	HIV; KO2011; ENVSBBH.			
DR	GLYCOSULEDB; P04582; .			
DR	InterPro; IPR000328; Env GP41.			
DR	InterPro; IPR000777; GP120.			
DR	Pfam; PF00516; GP120; 1.			
KW	AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal; 3D-structure.			
KW	SIGNAL	1	30	
FT	CHAIN	31	506	EXTERIOR MEMBRANE GLYCOPROTEIN.
FT	CHAIN	507	851	TRANSMEMBRANE GLYCOPROTEIN.
FT	DISULFID	54	74	BY SIMILARITY.
FT	DISULFID	119	205	BY SIMILARITY.
FT	DISULFID	126	196	BY SIMILARITY.
FT	DISULFID	131	157	BY SIMILARITY.
FT	DISULFID	218	247	BY SIMILARITY.
FT	DISULFID	228	239	BY SIMILARITY.
FT	DISULFID	296	331	BY SIMILARITY.
FT	DISULFID	378	440	BY SIMILARITY.

Query Match	85.4%;	Score 2861.5;	DB 1;	Length 851;
Best Local Similarity	83.8%;	Pred. No. 1.8e-220;		
Matches 542;	Conservative 33;	Mismatches 47;	Indels 25;	Gaps 5

Qy	2	KLWMTVYVGVVMEATTTLCASAPAKYDPEVHNVAHTHCVPDPDPOEVLNATVE	61
Db	32	EKLMTVIFGVFWMEATTTLCASDAPAKYDPEVHNVAHTHCVPDPDPOEVLNATVE	91
Qy	62	HNFMKNMNVBOMEDIIISLMDQSLKPCVKLTPLCVTLNCKDVANATNTNDSBG--TWBR	119
Db	92	NFMNKNMNVBOMEDIIISLMDQSLKPCVKLTPLCVTLNCKDVANATNTNDSBGRIMMK	151
Qy	120	GEIKKCSFNITTSIDDEVOKEXALFYKLDVXAI DNNNTSYRLISCDTSYITQACPKISFE	179
Db	152	GEIKKCSFNISSTSKGKQVEYAFPKLDIIIDINDTSYTLTSCNTSYITQACPKYSFE	211
Qy	180	PIPIHYCAPAGAILIKNDKTENGKXPCQCNVSTQCTHGIRPVSQTLLNGLSLAEVAV	239
Db	212	PIPIHYCAPAGAILIKNNKTENGCTGCPCTNVSTVQCTHGIRPVSQTLLNGLSLAEVAV	271
Qy	240	IRSDFTNNAKIIYOLKESVINEINCRPNPN-----GAGDIRQAHQ	280
Db	272	IRSVNFTDNAKIITVOLDTSVEINCTRPNNNTBKRIIRQGRGRARVTTIGKIGNMQAHQ	331
Qy	281	NISRAKANDTLKOIVIKLREOE-ENKTIIVFNHSSGGDPRIVMHSFNGGGEFFYCNSTOLF	339
Db	332	NISRAKMANATLKQIDSKLREQFGNNKTIIFKQSSGGDPRIVHVSFNGGGEFFYCNSTOLF	391
Qy	340	NSTWNNNTBGSNNTEG-NTITLPCRIKQIINNMOEYGMKAMVAPRINGQIRCSNITGLLL	398
Db	392	NSTW--STKSGNNTEGSDPITLPCRIKQIINNMOEYGMKAMVAPPIGQIRCSNITGLLL	449
Qy	399	TRDGGINENGTRIIFPGGGMDBMNRSELYKKVYVIEPLGVAPTKCKRKYVOREKRAVG	458
Db	450	TRDGGSNNSNESEIFRPGGGMDBMNRSELYKKVYVIEPLGVAPTAKRRVVOREKRAVG	509
Qy	459	IGAVNLFGLGAAGSTGGAASMTLTVQARLLLSGIVQOQNNLLRAIEAQORMLQITWAGIK	518
Db	510	IGALFLGFLGAAGSTGGAASMTLTVQARQLLSGIVQOQNNLLRAIEGQGHLLQITWAGIK	569
Qy	519	QLQAEVLAVERYLGDQQLIGWICSGKLLCCTAVPWNASWSNKSILDIRINNNTWMEBERE	578

Db	570	QLOARLTAERIKXQQLGIGWCGSKLICITPAFPMANSNSNLSLEGIYNNMTWME	629
Qy	579	IDNVTSEIYTLIEESONQOEKNEQELBELDYKASIMWFDTIKLWY	625
Db	630	INNYTSLHSLIEESONQOEKNEQELBELDYKASIMWFDTIKLWY	676
RESULT 15			
ENV_HV1W1	ENV_HV1W1	STANDARD;	PRT; 856 AA.
AC	P31872;		
DT	01-JUL-1993 (Rel. 26, Last sequence update)		
DT	01-JUL-1993 (Rel. 26, Last sequence update)		
DT	15-SEP-2003 (Rel. 42, Last annotation update)		
DE	Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].		
GN	ENV.		
OS	Human immunodeficiency virus type 1 (WM1 isolate) (HIV-1).		
OC	Viruses; Retroviral viruses; Retroviridae; Lentivirus.		
OX	NBI_TaxID=31678;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=86218077; PubMed=2423250;		
RA	Starcich B.R., Hann B.H., Shaw G.M., McNeely P.D., Modrow S., Wolf H., Parks E.S., Parks W.P., Josephs S.F., Gallo R.C., Wong-Staal F.;		
RA	"Identification and characterization of conserved and variable regions in the envelope gene of HTLV-III/LAV, the retrovirus of AIDS."		
RT			
RL	Cell 45:637-648 (1986).		
CC	-I- M5CGLANPEUS. ISOLATES WM1, WM2, AND WM3 WERE OBTAINED FROM BLOOD SAMPLES SEQUENTIALLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO WAS PERINATALLY INFECTED BY HER MOTHER.		
CC			
DR	PIR; A24774; VCLJ3W.		
DR	PDB; 1LB0; 04-DEC-02.		
DR	PDB; 1LCK; 04-DEC-02.		
DR	InterPro: IPR000328; Env GP41.		
DR	InterPro: IPR000777; GP120.		
DR	Pfam; Pf00516; GP120; 1.		
DR	Pfam; Pf00517; GP41; 1.		
KW	AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal; 3D-structure.		
FT	1	29	
FT	CHAIN	30	510
FT	CHAIN	511	856
FT	DISULFID	53	73
FT	DISULFID	118	205
FT	DISULFID	125	196
FT	DISULFID	130	152
FT	DISULFID	218	247
FT	DISULFID	228	239
FT	DISULFID	296	330
FT	DISULFID	376	444
FT	DISULFID	383	417
FT	CARBOHYD	87	87
FT	CARBOHYD	134	134
FT	CARBOHYD	140	140
FT	CARBOHYD	151	151
FT	CARBOHYD	155	155
FT	CARBOHYD	183	183
FT	CARBOHYD	197	197
FT	CARBOHYD	234	234
FT	CARBOHYD	241	241
FT	CARBOHYD	262	262
FT	CARBOHYD	276	276
FT	CARBOHYD	289	289
FT	CARBOHYD	295	295
FT	CARBOHYD	331	331
FT	CARBOHYD	338	338
FT	CARBOHYD	354	354
FT	CARBOHYD	360	360
FT	CARBOHYD	390	390
FT	CARBOHYD	394	394
FT	CHAIN	30	510
FT	CHAIN	511	856
FT	DISULFID	53	73
FT	DISULFID	118	205
FT	DISULFID	125	196
FT	DISULFID	130	152
FT	DISULFID	218	247
FT	DISULFID	228	239
FT	DISULFID	296	330
FT	DISULFID	376	444
FT	DISULFID	383	417
FT	CARBOHYD	87	87
FT	CARBOHYD	134	134
FT	CARBOHYD	140	140
FT	CARBOHYD	151	151
FT	CARBOHYD	155	155
FT	CARBOHYD	183	183
FT	CARBOHYD	197	197
FT	CARBOHYD	234	234
FT	CARBOHYD	241	241
FT	CARBOHYD	262	262
FT	CARBOHYD	276	276
FT	CARBOHYD	289	289
FT	CARBOHYD	295	295
FT	CARBOHYD	331	331
FT	CARBOHYD	338	338
FT	CARBOHYD	354	354
FT	CARBOHYD	360	360
FT	CARBOHYD	390	390
FT	CARBOHYD	394	394
FT	CHAIN	30	510
FT	CHAIN	511	856
FT	DISULFID	53	73
FT	DISULFID	118	205
FT	DISULFID	125	196
FT	DISULFID	130	152
FT	DISULFID	218	247
FT	DISULFID	228	239
FT	DISULFID	296	330
FT	DISULFID	376	444
FT	DISULFID	383	417
FT	CARBOHYD	87	87
FT	CARBOHYD	134	134
FT	CARBOHYD	140	140
FT	CARBOHYD	151	151
FT	CARBOHYD	155	155
FT	CARBOHYD	183	183
FT	CARBOHYD	197	197
FT	CARBOHYD	234	234
FT	CARBOHYD	241	241
FT	CARBOHYD	262	262
FT	CARBOHYD	276	276
FT	CARBOHYD	289	289
FT	CARBOHYD	295	295
FT	CARBOHYD	331	331
FT	CARBOHYD	338	338
FT	CARBOHYD	354	354
FT	CARBOHYD	360	360
FT	CARBOHYD	390	390
FT	CARBOHYD	394	394
FT	CHAIN	30	510
FT	CHAIN	511	856
FT	DISULFID	53	73
FT	DISULFID	118	205
FT	DISULFID	125	196
FT	DISULFID	130	152
FT	DISULFID	218	247
FT	DISULFID	228	239
FT	DISULFID	296	330
FT	DISULFID	376	444
FT	DISULFID	383	417
FT	CARBOHYD	87	87
FT	CARBOHYD	134	134
FT	CARBOHYD	140	140
FT	CARBOHYD	151	151
FT	CARBOHYD	155	155
FT	CARBOHYD	183	183
FT	CARBOHYD	197	197
FT	CARBOHYD	234	234
FT	CARBOHYD	241	241
FT	CARBOHYD	262	262
FT	CARBOHYD	276	276
FT	CARBOHYD	289	289
FT	CARBOHYD	295	295
FT	CARBOHYD	331	331
FT	CARBOHYD	338	338
FT	CARBOHYD	354	354
FT	CARBOHYD	360	360
FT	CARBOHYD	390	390
FT	CARBOHYD	394	394
FT	CHAIN	30	510
FT	CHAIN	511	856
FT	DISULFID	53	73
FT	DISULFID	118	205
FT	DISULFID	125	196
FT	DISULFID	130	152
FT	DISULFID	218	247
FT	DISULFID	228	239
FT	DISULFID	296	330
FT	DISULFID	376	444
FT	DISULFID	383	417
FT	CARBOHYD	87	87
FT	CARBOHYD	134	134
FT	CARBOHYD	140	140
FT	CARBOHYD	151	151
FT	CARBOHYD	155	155
FT	CARBOHYD	183	183
FT	CARBOHYD	197	197
FT	CARBOHYD	234	234
FT	CARBOHYD	241	241
FT	CARBOHYD	262	262
FT	CARBOHYD	276	276
FT	CARBOHYD	289	289
FT	CARBOHYD	295	295
FT	CARBOHYD	331	331
FT	CARBOHYD	338	338
FT	CARBOHYD	354	354
FT	CARBOHYD	360	360
FT	CARBOHYD	390	390
FT	CARBOHYD	394	394
FT	CHAIN	30	510
FT	CHAIN	511	856
FT	DISULFID	53	73
FT	DISULFID	118	205
FT	DISULFID	125	196
FT	DISULFID	130	152
FT	DISULFID	218	247
FT	DISULFID	228	239
FT	DISULFID	296	330
FT	DISULFID	376	444
FT	DISULFID	383	417
FT	CARBOHYD	87	87
FT	CARBOHYD	134	134
FT	CARBOHYD	140	140
FT	CARBOHYD	151	151
FT	CARBOHYD	155	155
FT	CARBOHYD	183	183
FT	CARBOHYD	197	197
FT	CARBOHYD	234	234
FT	CARBOHYD	241	241
FT	CARBOHYD	262	262
FT	CARBOHYD	276	276
FT	CARBOHYD	289	289
FT	CARBOHYD	295	295
FT	CARBOHYD	331	331
FT	CARBOHYD	338	338
FT	CARBOHYD	354	354
FT	CARBOHYD	360	360
FT	CARBOHYD	390	390
FT	CARBOHYD	394	394
FT	CHAIN	30	510
FT	CHAIN	511	856
FT	DISULFID	53	73
FT	DISULFID	118	205
FT	DISULFID	125	196
FT	DISULFID	130	152
FT	DISULFID	218	247
FT	DISULFID	228	239
FT	DISULFID	296	330
FT	DISULFID	376	444
FT	DISULFID	383	417
FT	CARBOHYD	87	87
FT	CARBOHYD	134	134
FT	CARBOHYD	140	140
FT	CARBOHYD	151	151
FT	CARBOHYD	155	155
FT	CARBOHYD	183	183
FT	CARBOHYD	197	197
FT	CARBOHYD	234	234
FT	CARBOHYD	241	241
FT	CARBOHYD	262	262
FT	CARBOHYD	276	276
FT	CARBOHYD	289	289
FT	CARBOHYD	295	295
FT	CARBOHYD	331	331
FT	CARBOHYD	338	338
FT	CARBOHYD	354	354
FT	CARBOHYD	360	360
FT	CARBOHYD	390	390
FT	CARBOHYD	394	394
FT	CHAIN	30	510
FT	CHAIN	511	856
FT	DISULFID	53	73
FT	DISULFID	118	205
FT	DISULFID	125	196
FT	DISULFID	130	152
FT	DISULFID	218	247
FT	DISULFID	228	239
FT	DISULFID	296	330
FT	DISULFID	376	444
FT	DISULFID	383	417
FT	CARBOHYD	87	87
FT	CARBOHYD	134	134
FT	CARBOHYD	140	140
FT	CARBOHYD	151	151
FT	CARBOHYD	155	155
FT	CARBOHYD	183	183
FT	CARBOHYD	197	197
FT	CARBOHYD	234	234
FT	CARBOHYD	241	241
FT	CARBOHYD	262	262
FT	CARBOHYD	276	276
FT	CARBOHYD	289	289
FT	CARBOHYD	295	295
FT	CARBOHYD	331	331
FT	CARBOHYD	338	338
FT	CARBOHYD	354	354
FT	CARBOHYD	360	360
FT	CARBOHYD	390	390
FT	CARBOHYD	394	394
FT	CHAIN	30	510
FT	CHAIN	511	856
FT	DISULFID	53	73
FT	DISULFID	118	205
FT	DISULFID	125	196
FT	DISULFID	130	152
FT	DISULFID	218	247
FT	DISULFID	228	239
FT	DISULFID	296	330
FT	DISULFID	376	444
FT	DISULFID	383	417
FT	CARBOHYD	87	87
FT	CARBOHYD	134	134
FT	CARBOHYD	140	140
FT	CARBOHYD	151	151
FT	CARBOHYD	155	155
FT	CARBOHYD	183	183
FT	CARBOHYD	197	197
FT	CARBOHYD	234	234
FT	CARBOHYD	241	241
FT	CARBOHYD	262	262
FT	CARBOHYD	276	276
FT	CARBOHYD	289	289
FT	CARBOHYD	295	295
FT	CARBOHYD	331	331
FT	CARBOHYD	338	338
FT	CARBOHYD	354	354
FT	CARBOHYD	360	360
FT	CARBOHYD	390	390
FT	CARBOHYD	394	394
FT	CHAIN	30	510
FT	CHAIN	511	856
FT	DISULFID	53	73
FT	DISULFID	118	205
FT	DISULFID	125	196
FT	DISULFID	130	152
FT	DISULFID	218	247
FT	DISULFID	228	239
FT	DISULFID	296	330
FT	DISULFID	376	444
FT	DISULFID	383	417
FT	CARBOHYD	87	87
FT	CARBOHYD	134	134
FT	CARBOHYD	140	140
FT	CARBOHYD	151	151
FT	CARBOHYD	155	155
FT	CARBOHYD	183	183
FT	CARBOHYD	197	197
FT	CARBOHYD	234	234
FT	CARBOHYD	241	241
FT	CARBOHYD	262	262
FT	CARBOHYD	276	276
FT	CARBOHYD	289	289
FT	CARBOHYD	295	295
FT	CARBOHYD	331	331
FT	CARBOHYD	338	338
FT	CARBOHYD	354	354
FT	CARBOHYD	360	360
FT	CARBOHYD	390	390
FT	CARBOHYD	394	394
FT	CHAIN	30	510
FT	CHAIN	511	856
FT	DISULFID	53	73
FT	DISULFID	118	205
FT	DISULFID	125	196
FT	DISULFID	130	152
FT	DISULFID	218	247
FT	DISULFID	228	239
FT	DISULFID	296	330
FT	DISULFID	376	444
FT	DISULFID	383	417
FT	CARBOHYD	87	87
FT	CARBOHYD	134	134
FT	CARBOHYD	140	140
FT	CARBOHYD	151	151
FT	CARBOHYD	155	155
FT	CARBOHYD	183	183
FT	CARBOHYD	197	197
FT	CARBOHYD	234	234
FT	CARBOHYD	241	

FT CARBOHYD 404 404 N-LINKED (GLCNAC. . .) (POTENTIAL) .
 FT CARBOHYD 447 N-LINKED (GLCNAC. . .) (POTENTIAL) .
 FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL) .
 FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL) .
 FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL) .
 FT CARBOHYD 625 625 N-LINKED (GLCNAC. . .) (POTENTIAL) .
 FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL) .
 SQ SEQUENCE 856 AA: 97526 MW: DB68D1B9C404DE9 CRC64;

Query Match 85.4%; Score 2861.5; DB 1; Length 856;
 Best Local Similarity 83.6%; Pred. No. 1.8e-220;
 Matches 547; Conservative 29; Mismatches 47; Indels 31; Gaps 6;

QY 1 VEKLVTVYVYGVVWKEATTTTLFCASDAKAYDTEVHNWATHACVPTDNPQEVLENT 60
 DB 30 VEQLVTVYVYGVVWKEATTTTLFCASDAKAYSTEAKWATHACVPTDNPQEVLENT 89
 QY 61 EHFNMKNMNVQWQEDIIISLWDSIKPCVKLTPLCVTLNCKDVANATNTTNDSEGTMERG 120
 DB 90 EHFNMKNMNVQWQEDIIISLWDSIKPCVKLTPLCVTLNCKDVANATNTTNDSEGTMERG 147
 QY 121 EIKNCSFNITTSIRDEQKEVALFKLDVYXIDNNN-----TSYRLISCDTSVITQACPK 175
 DB 148 EYKNCSEFNITTSIRKRVKEXVALFKLDVVPKSNNDSTYRRLHCHNTSVITQACSK 207
 QY 176 ISFEPPIHYCAPAGFALIKCNDKTFNGCKPCKNVSTXOCTHGIRPVVSTQLLNGSLAE 235
 DB 208 VSFEPPIHYCAPAGFALIKCNDKTFNGCKPCKNVSTVQCTHGIRPVVSTQLLNGSLAE 267
 QY 236 EEVVRSDNFTNNAKTIIVQLKESVEINCTRPNNNG-----AGDIRQ 277
 DB 268 EEVVRSENFDTNAKTIIVHLNESVEINCTRPNNVRRHHIHGPRAFYTGIEIRNIRQ 327
 QY 278 AHCNISRAKMDTLKOIVKLEOPENKTIIVNHSNGDPEIWMHSFNGGSEFFYCNSTQ 337
 DB 328 AHCNISRAKMDTLKOIVKLEOPENKTIIVNHSNGDPEIWMHSFNGGSEFFYCNSTQ 387
 QY 338 LFNSTWN--NNTGSSNTE--GNTITTLPCRICKIINNMQEVGKAMVAPPIRGQIRCSN 392
 DB 388 LFNSTWNVTGISTEONNTEENGDTTLPCRICKIINNMQEVGKAMVAPPIRGQIRCSN 447
 QY 393 ITGLLLTRDGINENGTEIFRPQGDMDRDNWSELYKYVVKIEPLGVAPTYCKRRVYQR 452
 DB 448 ITGLLLTRDGGNSSREEIFRPQGNMRDNWSELYKYVVKIEPLGVAPTYCKRRVYQR 507
 QY 453 EKRANG-IGAVFLGLGAAGSTMGASMTLTVQARLLSGIYQOQNNLLRAIEAQQRMQ 511
 DB 508 EKRANGAIGAMFLGLGAAGSTMGASMTLTVQARQLSGIYQOQNNLLRAIEAQQHMQ 567
 QY 512 LTVWGIKOLQARVLAVERYLADPQQLLGITWCGSGKLIICCTAVPMNAMSNSKSLDRIMNMT 571
 DB 568 LTVWGIKOLQARVLAVERYLADPQQLLGITWCGSGKLIICCTAVPMNAMSNSKSLDRIMNMT 627
 QY 572 WMEMEREIDNTYSEIYTLIEESQOQEKNOELBLEDKWASIMNPFDTIKMLMY 625
 DB 628 WMEMEREIDNTYSLIYNLIEESQOQEKNEQELLEDKWASIMNPFSTIMLMY 681

Search completed: December 12, 2003, 12:30:50
 Job time : 12.49 secs

THIS PAGE BLANK (08-70)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: December 12, 2003, 12:27:29 ; Search time 14.5506 Seconds
(without alignments)
4130.780 Million cell updates/sec

Title: US-10-032-162-17

Perfect score: 3351

Sequence: 1 VEKLMVTYVYGVPMWEKAT.....ELDKWASIMNFDITKMLWY 625

Scoring table: BLOSUM62

Searched: 283308 seqs, 96168682 residues

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: PIR_76:*

1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3293	98.3	847	2 T09448	envelope glycoprot
2	3264	97.4	847	2 S13289	env protein - huma
3	2901.5	86.6	861	1 VCLJLV	env polypeptide pr
4	2900.5	86.6	852	1 T12016	envelope glycoprot
5	2887	86.2	856	1 VCLJH3	env polypeptide pr
6	2883	86.0	856	1 VCLJVL	env polypeptide pr
7	2882	86.0	843	1 H44001	env polypeptide pr
8	2881	86.0	854	2 S13288	env protein - huma
9	2875	85.8	852	1 VCLJBR	env polypeptide -
10	2866.5	85.5	851	2 S33985	env polypeptide -
11	2861.5	85.4	856	1 VCLJ3W	env polypeptide pr
12	2859.5	85.3	859	1 VCLJMN	env polypeptide pr
13	2853	85.1	861	1 VCLJSC	env polypeptide pr
14	2836	84.6	855	1 VCLJSC	env polypeptide pr
15	2793.5	83.4	868	1 VCLJH4	env polypeptide -
16	2790.5	83.3	729	1 VCLJXK	env polypeptide pr
17	2790.5	83.3	729	1 VCLJXK	env polypeptide pr
18	2645	78.9	855	1 VCLJZR	env polypeptide pr
19	2636	78.7	853	2 S54384	envelope glycoprote
20	2578.5	76.9	846	1 VCLJND	env polypeptide pr
21	2572.5	76.8	859	2 T01672	envelope glycoprote
22	2517.5	75.1	856	1 A44963	env polypeptide pr
23	2180	65.1	854	1 VCLJST	env polypeptide pr
24	2058	61.4	506	2 A40218	envelop glycoprote
25	1890.5	56.4	443	2 C41621	env polypeptide P
26	1877.5	56.0	445	2 A41621	env polypeptide M
27	1865.5	55.7	495	2 S31493	env polypeptide -
28	1840	54.9	877	2 S49197	envelope protein P
29	1801	53.7	863	2 A53034	gag polypeptide -

30	1784	53.2	454	2 B41621	env polypeptide D
31	1210.5	36.1	290	2 S25940	env protein - huma
32	1204.5	35.9	297	2 S60538	envelope polypeptide
33	1195.5	35.7	852	1 VCLJG5	env polypeptide pr
34	1195	35.7	859	1 VCLJST	env polypeptide pr
35	1176	35.1	712	1 VCLJH4	env polypeptide pr
36	1165	34.8	877	2 C46356	env polypeptide pr
37	1162	34.7	858	1 VCLJG2	env polypeptide pr
38	1158	34.6	864	1 VCLJG4	env polypeptide pr
39	1148.5	34.3	859	1 VCLJCT	env polypeptide pr
40	1138	34.0	732	2 S46352	env polypeptide -
41	1136	33.9	881	1 VCLJG3	env polypeptide -
42	1135.5	33.9	851	2 S12159	env protein - huma
43	1135.5	33.9	869	2 A47665	env protein gp120 (
44	1134.5	33.9	859	2 S24571	env protein - huma
45	1134	33.8	869	2 S53098	envelope polypeptide

ALIGNMENTS

QY	1	VEKLMVTYVYGVPMWEKATTTLCASDAKADTEVHNVATHACVPTDPNPOEVLEANT	60
DB	30	VEKLMVTYVYGVPMWEKATTTLCASDAKADTEVHNVATHACVPTDPNPOEVLEANT	89
QY	61	EHFMKNNMVEQOEDIIISLMDSLKPCVLTPLCVTLNCKDVATNTNDSGTMERG	120
DB	90	EHFMKNNMVEQOEDIIISLMDSLKPCVLTPLCVTLNCKDVATNTNDSGTMERG	149
QY	121	EIKKCSFNTTTSIDDEVQKVALFYKLDVYKIDNNNTSYRLISCTSYITQACPKISSEP	180
DB	150	EIKKCSFNTTTSIDDEVQKVALFYKLDVYKIDNNNTSYRLISCTSYITQACPKISSEP	209
QY	181	IPHYCAPAGFAIILKNDKXTNGKPKCNVSTXCTGHSIRRVSTOLLNSLAEVVI	240
DB	210	IPHYCAPAGFAIILKNDKXTNGKPKCNVSTXCTGHSIRRVSTOLLNSLAEVVI	269
QY	241	RSDFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRQAHGNI	282
DB	270	RSDFTNNAKTIIVOLKESVEINCTRPNNNRKSHIGRAFTYTGSHIIDIRAHGNI	329
QY	283	SPRAKNDTLKOIVIKLRQFENKTIIVNHSGGDEIYVMSFGCGEFPYCNSTOLFNST	342
DB	330	SPRAKNDTLKOIVIKLRQFENKTIIVNHSFGCGDEIYVMSFGCGEFPYCNSTOLFNST	389
QY	343	WNNTEGSGNTEGNTITLPCRIKQIINMVGKAMVAPPIRGQIRGCSNTTGLLTGDS	402
DB	390	WNNTEGSGNTEGNTITLPCRIKQIINMVGKAMVAPPIRGQIRGCSNTTGLLTGDS	449
QY	403	GINENGTEIFRPGGDMKDNMRSELYKKYKVKIIEPLGVAPTKCKRRVVRKRAVGIGAV	462

Db 450 GINENGTEIFRPGGDMRDNRSELYKKVYKIEPLGVAFTKAKRVRVQREKRAVGIGAV 509
 Qy 463 FLGLGAGSTMGASMTLTVOARLLLSGIYQOONNLLRAIEAOORMLQLTWAGIKOLA 522
 Db 510 FLGLGAGSTMGASMTLTVOARLLLSGIYQOONNLLRAIEAOORMLQLTWAGIKOLA 569
 Qy 523 RVAVERYLGDQOOLGIGWCGSKLICTTAVPMNASWSNKSIDRIIMNMWEMEREIDNY 582
 Db 570 RVAVERYLGDQOOLGIGWCGSKLICTTAVPMNASWSNKSIDRIIMNMWEMEREIDNY 629
 Qy 583 TSEIYTLIESONOEKNEOELLELDKASIMNFDITKMLMY 625
 Db 630 TSEIYTLIESONOEKNEOELLELDKASIMNFDITKMLMY 672

RESULT 2

env protein - human immunodeficiency virus type 1
 C:Species: human immunodeficiency virus type 1, HIV-1
 Date: 19-Mar-1997 #sequence_reviseion 19-Mar-1997 #text_change 31-Oct-1997
 Accession: S13289
 R:O'Brien, W.A.; Koyanagi, Y.; Namazie, A.; Zhao, J.Q.; Diagne, A.; Idler, K.; Zack, J.A.
 Nature 348, 69-73, 1990
 A:Title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120
 A:Reference number: S13288; MUID:91043044; PMID:2172833
 A:Accession: S13289
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-847 <OBR>
 C:Superfamily: type E retrovirus env polyprotein

Query Match 97.4%; Score 3264; DB 2; Length 847;
 Best Local Similarity 95.5%; Pred. No. 7,2e-235;

Matches 613; Conservative 1; Mismatches 10; Indels 18; Gaps 1;

Qy 2 EKLWTVYYGVVWKKEATTLTFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNVTE 61
 Db 31 EKLWTVYYGVVWKKEATTLTFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNVTE 90
 Qy 62 HNNWKNMWEQMOEDIIISLDOSLKPCKVLTPLCVTLNCKDVNAITVNDSEGTMERGE 121
 Db 91 HNNWKNMWEQMOEDIIISLDOSLKPCKVLTPLCVTLNCKDVNAITVNDSEGTMERGE 150
 Qy 122 IKNSFNITTSIRDEVOKEVALFKLDVYKIDNNNTSRYLSCTPSVTTOACPKISPEPI 181
 Db 151 IKNSFNITTSIRDEVOKEVALFKLDVYKIDNNNTSRYLSCTPSVTTOACPKISPEPI 210
 Qy 182 PIHYCAPAGFAILKCNKTFNGKXPCXNVSTXCTGIRPVVSTOLLNGLAEVEVIR 241
 Db 211 PIHYCAPAGFAILKCNKTFNGKXPCXNVSTXCTGIRPVVSTOLLNGLAEVEVIR 270
 Qy 242 SDNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRAHCHIS 283
 Db 271 SDNFTNNAKTIIVOLKESVEINCTRPNNNRKSIHIGBFAFYTGELIGDIRAHCHIS 330
 Qy 284 RAKNNDTLKOIVILRBOFEKTIIVFNHSSGDEPIVHNSFCGSEFFYCNSTOLFENSTW 343
 Db 331 RAKNNDTLKOIVILRBOFEKTIIVFNHSSGDEPIVHNSFCGSEFFYCNSTOLFENSTW 390
 Qy 344 NNNTGSGNNTGNTTILPCRIKOIIMWQEVGKAMVAPPIRGQIRCSNITGLLTRDG 403
 Db 391 NNNTGSGNNTGNTTILPCRIKOIIMWQEVGKAMVAPPIRGQIRCSNITGLLTRDG 450
 Qy 404 INENGTETFRPGGDMRDNRSELYKKVYKIEPLGVAFTKAKRVRVQREKRAVGIGAV 463
 Db 451 INENGTETFRPGGDMRDNRSELYKKVYKIEPLGVAFTKAKRVRVQREKRAVGIGAV 510
 Qy 464 LGFLGAGSTMGASMTLTVOARLLLSGIYQOONNLLRAIEAOORMLQLTWAGIKOLA 523
 Db 511 LGFLGAGSTMGASMTLTVOARLLLSGIYQOONNLLRAIEAOORMLQLTWAGIKOLA 570
 Qy 524 VLAVERYLGDQOOLGIGWCGSKLICTTAVPMNASWSNKSIDRIIMNMWEMEREIDNY 583

Db 571 VLAVERYLGDQOOLGIGWCGSKLICTTAVPMNASWSNKSIDRIIMNMWEMEREIDNY 630
 Qy 584 SEIYTLIESONOEKNEOELLELDKASIMNFDITKMLMY 625
 Db 631 SEIYTLIESONOEKNEOELLELDKASIMNFDITKMLMY 672

RESULT 3

env polyprotein precursor - human immunodeficiency virus type 1 (isolate LAV-1a)

N:Alternate names: coat polyprotein

C:Species: human immunodeficiency virus type 1, HIV-1

A:Note: host Homo sapiens (man)

C:Date: 17-May-1985 #sequence_reviseion 17-May-1985 #text_change 16-Jul-1999

C:Accession: A03975

R:Main-Hobson, S.; Sonigo, P.; Danos, O.; Cole, S.; Alizon, M.

Cell 40, 9-17, 1985

A:Title: Nucleotide sequence of the AIDS virus, LAV.

A:Reference number: A90866; MUID:8509333; PMID:2981635

A:Accession: A03975

A:Molecule type: DNA

A:Residues: 1-861 <NAI>

A:Cross-References: GB:K02013; NID:G326417; PIDN:AAB59751.1; PID:G326424

C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprot

C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprot

F:1-30/Domain: signal sequence #status predicted <SIG>

F:31-516/Product: exterior membrane glycoprotein #status predicted <EX>

F:517-861/Product: transmembrane glycoprotein #status predicted <TM>

F:86,136,141,146,161,165,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411,

F:616,621,630,642,679,755,821/Binding site: carbohydrate (Asn) (covalent) #status predict

Query Match 86.6%; Score 2901.5; DB 1; Length 861;
 Best Local Similarity 84.3%; Pred. No. 7,2e-208;

Matches 552; Conservative 31; Mismatches 41; Indels 31; Gaps 7;

Qy 2 EKLWTVYYGVVWKKEATTLTFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNVTE 61
 Db 32 EKLWTVYYGVVWKKEATTLTFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNVTE 91
 Qy 62 HNNWKNMWEQMOEDIIISLDOSLKPCKVLTPLCVTLNCKDVNAITVNDSEGTMERGE 115
 Db 92 HNNWKNMWEQMOEDIIISLDOSLKPCKVLTPLCVTLNCKDVNAITVNDSEGTMERGE 151
 Qy 116 -TMRGEEKKCSFNITTSIRDEVOKEVALFKLDVYKIDNNNTSRYLSCTPSVTTOACPK 174
 Db 152 -TMRGEEKKCSFNITTSIRDEVOKEVALFKLDVYKIDNNNTSRYLSCTPSVTTOACPK 211
 Qy 175 KISPEPIHYCAPAGFAILKCNKTFNGKXPCXNVSTXCTGIRPVVSTOLLNGLAEVEVIR 234
 Db 212 KISPEPIHYCAPAGFAILKCNKTFNGKXPCXNVSTXCTGIRPVVSTOLLNGLAEVEVIR 271
 Qy 235 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----GAGDI 275
 Db 272 EEEVIRSDNFTNNAKTIIVOLKESVEINCTRPNNNRKSIHIGBFAFYTGELIGDIRAHCHIS 331
 Qy 276 RAKNNDTLKOIVILRBOFEKTIIVFNHSSGDEPIVHNSFCGSEFFYCNSTOLFENSTW 334
 Db 332 RAKNNDTLKOIVILRBOFEKTIIVFNHSSGDEPIVHNSFCGSEFFYCNSTOLFENSTW 391
 Qy 335 STOLFENSTWNTSMTSGSNNTGSDITTLPCRIKOIIMWQEVGKAMVAPPIRGQIRCS 390
 Db 392 STOLFENSTWNTSMTSGSNNTGSDITTLPCRIKOIIMWQEVGKAMVAPPIRGQIRCS 451
 Qy 391 SNITGLLTRDGGINENGTEIFRPGGDMRDNRSELYKKVYKIEPLGVAFTKAKRVRV 450
 Db 452 SNITGLLTRDGGINENGTEIFRPGGDMRDNRSELYKKVYKIEPLGVAFTKAKRVRV 511
 Qy 451 QREKRAVGIGAVFLGLGAGSTMGASMTLTVOARLLLSGIYQOONNLLRAIEAOORML 510
 Db 512 QREKRAVGIGAVFLGLGAGSTMGASMTLTVOARLLLSGIYQOONNLLRAIEAOORML 571

QY 511 QLTWGTGICQOLARVLAVERYLGDQQLIGWCGSKLICCTTAVPWNASWSNKSIDRIWMNM 570
 Db 572 QLTWGTGICQOLARVLAVERYLGDQQLIGWCGSKLICCTTAVPWNASWSNKSIDRIWMNM 631
 QY 571 TMMWERIDNTSYTLIESONQOEKNEBELLELDKMASLWMPDITKMLWY 625
 Db 632 TMMWERIDNTSYTLIESONQOEKNEBELLELDKMASLWMPDITKMLWY 666

RESULT 4

T12016

envelope glycoprotein - human immunodeficiency virus type 1 (strain sc14.3)

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000

C:Accession: T12016

R:McCutchan, F.E.; Sanders-Buell, E.; Salminen, M.O.; Carr, J.K.; Sheppard, W.H.

AIDS Res. Hum. Retroviruses 14, 329-337, 1998

A:Title: Diversity of the human immunodeficiency virus type 1 envelope glycoprotein in S

A:Reference number: 217379; PMID:98178716; PMID:9518994

Accession: T12016

Status: preliminary; translated from GB/EMBL/DBJ

Molecule type: DNA

A:Residues: 1-852 <MC>

A:Cross-references: EMBL:U90934; NID:g2351783; PIDN:AACS9271.1; PID:g2351784

C:Genetics:

A:Gene: env

C:Superfamily: type E retrovirus env polypeptide

Query Match 86.6%; Score 2900.5; DB 2; Length 852;

Best Local Similarity 85.0%; Pred. No. 8.5e-208;

Matches 550; Conservative 30; Mismatches 44; Indels 23; Gaps 5;

QY 2 EKLWTVYVYGVPMKEATTTLFCASDAKAYDTEVNVNATHACVPTDNPQEVLENTYE 61
 Db 31 EQLRTVYVYGVPMKEATTTLFCASDAKAYDTEVNVNATHACVPTDNPQEVLENTYE 90
 QY 62 HNNMKNMNMVQMOEDIIISLDQSLKPCVKLTPLCVTLNCKDV--NATVTTNDSGTEWR 119
 Db 91 NFNMMKNMNMVQMOEDIIISLDQSLKPCVKLTPLCVTLNCKDV--NATVTTNDSGTEWR 150
 QY 120 GEIKNCSFNITTSIRDEYQKVALFYKLDVYXIDNNNTSYRLISGDTSVITQACPKISFE 179
 Db 151 GEIKNCSFNITTSIRKQKVALFYKLDVYXIDNNNTSYRLISGDTSVITQACPKISFE 210
 QY 180 PIPHYCAPAGFALIKCNDKTFNGKPCKNVSTXOCTGIRPVYSTOLLNGLAEEVY 239
 Db 211 PIPHYCAPAGFALIKCNDKTFNGKPCKNVSTXOCTGIRPVYSTOLLNGLAEEVY 270
 QY 240 IRSDFNTNNAKTIIVQLKESVEINCTRPNN-----GA--GDIRQAHN 281
 Db 271 IRSEFTNNAKTIIVQLKESVEINCTRPNN-----GA--GDIRQAHN 330
 QY 282 ISRAKMNNTLKOIYKLBQFENKTIIVNHSFGDEIVTHSHFNGCGEFPYCNSTOLF 341
 Db 331 ISRAKMNNTLKOIYKLBQFENKTIIVNHSFGDEIVTHSHFNGCGEFPYCNSTOLF 390
 QY 342 TM--NNTEGSNNTEG--TITLPCRIOIINMOWGVGAMVAPPIRGOIRCSNNTG 398
 Db 391 TMMFNNTNNDTEGNGTITLPCRIOIINMOWGVGAMVAPPIRGOIRCSNNTG 450
 QY 399 TRDGINENGEIIFRPGGDMRDMNRSELKYKVKIIEPLGVAPTKCRKRVVQREKRAVG 458
 Db 451 TRDGTNNSTMTETFRPGGDMRDMNRSELKYKVKIIEPLGVAPTKCRKRVVQREKRAVG 510
 QY 459 IGAFLGLFLGAAGSTMGAASTLTLYQARLLLSGIVQOONNLRAIEAOQRMQLTWG 518
 Db 511 IGAFLGLFLGAAGSTMGAASTLTLYQARLLLSGIVQOONNLRAIEAOQRMQLTWG 570
 QY 519 QLOARVLAVERYLGDQQLIGWCGSKLICCTTAVPWNASWSNKSIDRIWMNM 578
 Db 571 QLOARVLAVERYLGDQQLIGWCGSKLICCTTAVPWNASWSNKSIDRIWMNM 630
 QY 579 IDNTSYTLIESONQOEKNEBELLELDKMASLWMPDITKMLWY 625

Db 631 IDNTSYTLIESONQOEKNEBELLELDKMASLWMPDITKMLWY 677

RESULT 5

VCLH3

env polypeptide precursor - human immunodeficiency virus type 1 (isolate HTLV-III, BH10)

N:Alternate names: coat polypeptide

C:Species: human immunodeficiency virus type 1, HIV-1

A:Note: host Homo sapiens (man)

C:Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 16-Jul-1999

C:Accession: A03973

R:Ratner, L.; Haseltine, W.; Patarca, R.; Livak, K.J.; Starcich, B.; Josephs, S.F.; Doran

berger, J.A.; Papas, T.S.; Chirayeb, J.; Chang, N.T.; Gallo, R.C.; Wong-Staal, F.

Nature 313, 277-284, 1985

A:Title: Complete nucleotide sequence of the AIDS virus, HTLV-III.

A:Reference number: A93353; PMID:85111123; PMID:2578615

Accession: A03973

Molecule type: DNA

A:Residues: 1-856 <RAT>

A:Cross-references: GB:M15654; GB:K02008; GB:K02009; GB:K02010; NID:g326383; PIDN:AAA442

C:Genetics:

A:Gene: env

C:Superfamily: type E retrovirus env polypeptide

C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polypot

F.1-30/Domains: signal sequence #status predicted <SIG>

F.31-51/Product: exterior membrane glycoprotein #status predicted <EXT>

F.512-856/Product: transmembrane glycoprotein #status predicted <TM>

F.188,136,141,156,160,186,197,230,234,241,262,276,289,295,301,332,339,356,386,392,397,406

F.611,616,625,637,674,750,816/Binding site: carbohydrate (Asn) (covalent) #status predic

Query Match 86.2%; Score 2887; DB 1; Length 856;
 Best Local Similarity 84.0%; Pred. No. 8.6e-207;
 Matches 546; Conservative 34; Mismatches 44; Indels 26; Gaps 5;

QY 2 EKLWTVYVYGVPMKEATTTLFCASDAKAYDTEVNVNATHACVPTDNPQEVLENTYE 61
 Db 32 EKLWTVYVYGVPMKEATTTLFCASDAKAYDTEVNVNATHACVPTDNPQEVLENTYE 91
 QY 62 HNNMKNMNMVQMOEDIIISLDQSLKPCVKLTPLCVTLNCKDV--NATVTTNDSGTEWR 119
 Db 92 NFNMMKNMNMVQMOEDIIISLDQSLKPCVKLTPLCVTLNCKDV--NATVTTNDSGTEWR 151
 QY 120 GEIKNCSFNITTSIRDEYQKVALFYKLDVYXIDNNNTSYRLISGDTSVITQACPKISFE 179
 Db 152 GEIKNCSFNITTSIRKQKVALFYKLDVYXIDNNNTSYRLISGDTSVITQACPKISFE 211
 QY 180 PIPHYCAPAGFALIKCNDKTFNGKPCKNVSTXOCTGIRPVYSTOLLNGLAEEVY 239
 Db 212 PIPHYCAPAGFALIKCNDKTFNGKPCKNVSTXOCTGIRPVYSTOLLNGLAEEVY 271
 QY 240 IRSDFNTNNAKTIIVQLKESVEINCTRPNN-----GA--GDIRQAHN 280
 Db 272 IRSEFTNNAKTIIVQLKESVEINCTRPNN-----GA--GDIRQAHN 331
 QY 282 ISRAKMNNTLKOIYKLBQFENKTIIVNHSFGDEIVTHSHFNGCGEFPYCNSTOLF 339
 Db 332 ISRAKMNNTLKOIYKLBQFENKTIIVNHSFGDEIVTHSHFNGCGEFPYCNSTOLF 391
 QY 340 NSTWNN--NTEGSNNTEG--TITLPCRIOIINMOWGVGAMVAPPIRGOIRCSNNTG 395
 Db 392 NSTWNN--NTEGSNNTEG--TITLPCRIOIINMOWGVGAMVAPPIRGOIRCSNNTG 451
 QY 396 LLLTRDGINENGEIIFRPGGDMRDMNRSELKYKVKIIEPLGVAPTKCRKRVVQREKRAVG 455
 Db 452 LLLTRDGINENGEIIFRPGGDMRDMNRSELKYKVKIIEPLGVAPTKCRKRVVQREKRAVG 511
 QY 456 IGAFLGLFLGAAGSTMGAASTLTLYQARLLLSGIVQOONNLRAIEAOQRMQLTWG 515
 Db 512 IGAFLGLFLGAAGSTMGAASTLTLYQARLLLSGIVQOONNLRAIEAOQRMQLTWG 571
 QY 516 QLOARVLAVERYLGDQQLIGWCGSKLICCTTAVPWNASWSNKSIDRIWMNM 575

Db 572 GIKOLQARILAVERYLKDQQLIGWCGSKLICCTTAVPWNASWNSKSLBQIMNMWMEW 631
Qy 576 EREIDNTSEIYTLIESQNOEKNEBELLDKMASLWNPDTIKMLY 625
Db 632 DREINNTSLIHSILIESQNOEKNEBELLDKMASLWNPDTIKMLY 681

RESULT 6
VC/L/VL

env polyprotein precursor - human immunodeficiency virus type 1 (isolate LV)

N/Alternate names: coat polyprotein
C/Species: human immunodeficiency virus type 1, HIV-1
A/Note: host Homo sapiens (man)
C/Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 16-Jul-1999
C/Accession: A03974
R/Mesling, M.A.; Smith, D.H.; Cabrerilla, C.D.; Benton, C.V.; Leaky, L.A.; Capon, D.J.
Nature 313, 450-458, 1985
A/Title: Nucleic acid structure and expression of the human AIDS/lymphadenopathy retrovi
A/Reference number: A93355; MUID:8511157; PMID:2982104
C/Accession: A03974
Molecule type: DNA

A/Residues: 1-856 <MB>

A/Cross-references: GB:K02083; NID:G555008; PIDN:AB59873.1; PID:G328559

C/Genetics:

C/Superfamily: type E retrovirus env polyprotein

C/Keywords: AIDS: capsid protein; coat protein; glycoprotein; immunodeficiency; polypro

F:1-30/Domain: signal sequence #status predicted <SIG>

F:31-51/Domain: signal sequence #status predicted <SIG>

F:52-856/Product: transmembrane glycoprotein #status predicted <TM>

F:88,136,141,156,160,186,197,230,241,262,276,289,295,301,332,339,356,386,392,397,406

F:611,616,625,637,674,750,816/Binding site: carbohydrate (Asn) (covalent) #status predic

Query Match 86.0%; Score 2883; DB 1; Length 856;

Best Local Similarity 83.8%; Pred. No. 1.7e-206;

Matches 545; Conservative 33; Mismatches 46; Indels 26; Gaps 5;

Qy 2 EKLWTVYGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDPNPOEVLLENTE 61

Db 32 EKLWTVYGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDPNPOEVLLENTE 91

Qy 62 HFNWKNWWEQWQEDIIISLWDSLKPCVKLTPLCVTLNCKDVNATNTNDSG--TWER 119

Db 92 NFNWKNWWEQWQEDIIISLWDSLKPCVKLTPLCVTLNCKDVNATNTNDSG--TWER 151

Qy 120 GEIKNSFNITTSIRDEYQKAYALFYKLDVXIDNNNTSYRLISGDSVTIQAQPKISFE 179

Db 152 GEIKNSFNITTSIRDEYQKAYALFYKLDVXIDNNNTSYRLISGDSVTIQAQPKISFE 211

Qy 180 PIPHYCAPAGFALIKNDKTFNGKPCKNVSTXOCTGIRPVYSTOLLNGSLAEEBIV 239

Db 212 PIPHYCAPAGFALIKNDKTFNGKPCKNVSTXOCTGIRPVYSTOLLNGSLAEEBIV 271

Qy 240 IRSNFTNNAKTIIVOLKESVEINCTRPNN--GAGDIRQAH 280

Db 272 IRSNFTNNAKTIIVOLKESVEINCTRPNN--GAGDIRQAH 331

Qy 281 NISAKNNDTLKQIVIKRBOF-EKTIIVFNHSSGDEIYVHNSGGEFFYCNSDTLF 339

Db 332 NISAKNNDTLKQIVIKRBOF-EKTIIVFNHSSGDEIYVHNSGGEFFYCNSDTLF 391

Qy 340 NSTNN--NTEGSNTEG--NTITLPCRIKQIIMMOWEYKAMAPPIRGQIRCSNITG 395

Db 392 NSTNN--NTEGSNTEG--NTITLPCRIKQIIMMOWEYKAMAPPIRGQIRCSNITG 451

Qy 396 LLTLTRDGINENGEIIFRPGGDMRDMNRSELYKYVVKIEPLGVAPTKGRVVOREKR 455

Db 452 LLTLTRDGINENGEIIFRPGGDMRDMNRSELYKYVVKIEPLGVAPTKGRVVOREKR 511

Qy 456 AVGIGAVFLGLGAAGSTGAASMTLTQARLLLSGIYQOONNLIRAEAOQRMQLTWM 515

Db 512 AVGIGAVFLGLGAAGSTGAASMTLTQARLLLSGIYQOONNLIRAEAOQRMQLTWM 571

Qy 516 GIKOLQARILAVERYLKDQQLIGWCGSKLICCTTAVPWNASWNSKSLBQIMNMWMEW 575
Db 572 GIKOLQARILAVERYLKDQQLIGWCGSKLICCTTAVPWNASWNSKSLBQIMNMWMEW 631
Qy 576 EREIDNTSEIYTLIESQNOEKNEBELLDKMASLWNPDTIKMLY 625
Db 632 DREINNTSLIHSILIESQNOEKNEBELLDKMASLWNPDTIKMLY 681

RESULT 7

env polyprotein precursor - human immunodeficiency virus type 1 (strain YU-2)

N/Alternate names: coat polyprotein

C/Species: human immunodeficiency virus type 1, HIV-1

A/Note: host Homo sapiens (man)

C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Sep-1994

C/Accession: H44001

R/Li, Y.; Hui, H.; Burgess, C.J.; Price, R.W.; Sharp, P.M.; Hahn, B.H.; Shaw, G.M.

J. Virol. 66, 6587-6600, 1992

A/Title: Complete nucleotide sequence, genome organization, and biological properties of

A/Reference number: A44001; MUID:53021387; PMID:1404605

A/Accession: H44001

A/Residues: 1-843

A/Cross-references: GB:M93258

C/Genetics:

C/Superfamily: type E retrovirus env polyprotein

C/Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein

F:1-29/Domain: signal sequence #status predicted <SIG>

F:30-489/Product: coat protein gp120 #status predicted <GP1>

F:490-843/Product: coat protein gp41 #status predicted <GP2>

F:499-515/Region: hydrophobic

F:673-689/Region: hydrophobic

F:738-755/Domain: transmembrane #status predicted <TM>

F:87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

Query Match 86.0%; Score 2882; DB 1; Length 843;

Best Local Similarity 85.1%; Pred. No. 2e-206;

Matches 548; Conservative 28; Mismatches 42; Indels 26; Gaps 7;

Qy 2 EKLWTVYGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDPNPOEVLLENTE 61

Db 31 EKLWTVYGVVWKEATTTTFCASDAKAYDTEVHNWATHACVPTDPNPOEVLLENTE 90

Qy 62 HFNWKNWWEQWQEDIIISLWDSLKPCVKLTPLCVTLNCKDVNATNTNDSG--TWER 120

Db 91 NFNWKNWWEQWQEDIIISLWDSLKPCVKLTPLCVTLNCKDVNATNTNDSG--TWER 150

Qy 121 EIKKNSFNITTSIRDEYQKAYALFYKLDVXIDNNNTSYRLISGDSVTIQAQPKISFE 180

Db 151 EIKKNSFNITTSIRDEYQKAYALFYKLDVXIDNNNTSYRLISGDSVTIQAQPKISFE 208

Qy 181 IPIHYCAPAGFALIKNDKTFNGKPCKNVSTXOCTGIRPVYSTOLLNGSLAEEBIV 240

Db 209 IPIHYCAPAGFALIKNDKTFNGKPCKNVSTXOCTGIRPVYSTOLLNGSLAEEBIV 268

Qy 241 RSNFTNNAKTIIVOLKESVEINCTRPNN--GAGDIRQAH 282

Db 269 RSNFTNNAKTIIVOLKESVEINCTRPNN--GAGDIRQAH 328

Qy 283 SRKANDTLKQIVIKRBOF-EKTIIVFNHSSGDEIYVHNSGGEFFYCNSDTLF 341

Db 329 SKTQWETLLEDAIKLEQEGNNKTIIFNPSGSDPEIYVHNSGGEFFYCNSDTLF 386

Qy 342 TANNNTGSNNTEGNTITLPCRIKQIIMMOWEYKAMAPPIRGQIRCSNITGLTTRD 401

Db 387 TM-NDTRKLNVT--GRNTTLPCRIKQIIMMOWEYKAMAPPIRGQIRCSNITGLTTRD 444

Qy 402 GGINENGEIIFRPGGDMRDMNRSELYKYVVKIEPLGVAPTKGRVVOREKRAVIGA 461

Db 445 GGGKOTNGEIRFRGGGDMRDNMRSELYKKYKVKIEPLGVAPTKRRVYQREKRAVGLGA 504

Qy 462 VFLEGLGAAGSTMGASMTLTVOARLLLSGIYQOQNNLLRAIEAOQRMQLTWGIGKQO 521

Db 505 LFLGLGAGSTMGASMTLTVOARQLSGIYQOQNNLLRAIEAOQHLQLTWGIGKQO 564

Qy 522 ARVLAVERYLGDQQLLGIGWCSGSKLICCTAVPWNASWSNKSIDRIIMNNMTMWEEREIDN 581

Db 565 ARVLAVERYLRDQQLLGIGWCSGSKLICCTTPPWNMTWSNKSINELINDNTMWMKEREIDN 624

Qy 582 YTSSEYTLIEESQNOEKNEQBELBLDKASIMNFDITKMLY 625

Db 625 YTHIYSLIEESQNOEKNEQBELBLDKASIMNFDITKMLY 668

RESULT 8

env protein - human immunodeficiency virus type 1
 C:Species: human immunodeficiency virus type 1, HIV-1
 Accession: S13288
 Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997
 Accession: S13288
 A:Title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120
 A:Reference number: S13288; PMID:91043044; PMID:2172833
 A:Accession: S13288
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-854 <OBR>
 C:Superfamily: type E retrovirus env polypeptide

Query Match 86.0%; Score 2881; DB 2; Length 854;
 Best Local Similarity 84.2%; Pred. No. 2, 4e-206;

Matches 547; Conservative 29; Mismatches 46; Indels 28; Gaps 6;

Qy 2 EKLWTVYVGVVPMKEATTLTFCASDAKAYDTEVHNWATACVPTDNPQEVLENYTE 61

Db 32 EKLWTVYVGVVPMKEATTLTFCASDAKAYDTEVHNWATACVPTDNPQEVLENYTE 91

Qy 62 HFNMMKNMVMQOMODIISLDQSLKPCVKLTPLCTVLANCKDVNATNTTNSSEGE--THER 119

Db 92 HFNMMKNMVMQOMODIISLDQSLKPCVKLTPLCTVLANCKDVNATNTTNSSEGE--THER 151

Qy 120 GEIKKCSFNITTSIRDEVOKEVALFYKLDVYKIDNNNSYRLISCDTSVITQACPKISFE 179

Db 152 GEIKKCSFNITTSIRDEVOKEVALFYKLDVYKIDNNNSYRLISCDTSVITQACPKISFE 209

Qy 160 PIPHYCAPAGFAILKCKNDKTFNGKXPCKNVSTXOCTGIRPVVSTOLLNGSLAEEV 239

Db 210 PIPHYCAPAGFAILKCKNDKTFNGKXPCKNVSTXOCTGIRPVVSTOLLNGSLAEEV 269

Qy 240 IRSDNFTNNAKTIIVQLKESVEINCTRPNNN-----GAGDIRQAHN 280

Db 270 IRSDNFTNNAKTIIVQLKESVEINCTRPNNN-----GAGDIRQAHN 329

Qy 281 NISRAKNDTLKQIVYKLRQEF--ENKTIYFNSSGGDEIVMHSFNGGEPFYCNSTOLF 339

Db 330 NISRAKNDTLKQIVYKLRQEF--ENKTIYFNSSGGDEIVMHSFNGGEPFYCNSTOLF 389

Qy 340 NSTMNN--NTEGNNTEG--NTITLPCRKQIIMMOEVGKAMVAPPIRGOIRCSNITG 395

Db 390 NSTMNN--NTEGNNTEG--NTITLPCRKQIIMMOEVGKAMVAPPIRGOIRCSNITG 449

Qy 396 LLITLDGGINENGTIFRPGGDMRDNMRSELYKKYKVKIEPLGVAPTKRRVYQREK 455

Db 450 LLITLDGGINENGTIFRPGGDMRDNMRSELYKKYKVKIEPLGVAPTKRRVYQREK 509

Qy 456 AVGIGAVFLGAGSTMGASMTLTVOARLLLSGIYQOQNNLLRAIEAOQRMQLTWG 515

Db 510 AVGIGAVFLGAGSTMGASMTLTVOARQLSGIYQOQNNLLRAIEAOQHLQLTW 569

Qy 516 GIKQLQARVLAVERYLGDQQLLGIGWCSGSKLICCTAVPWNASWSNKSIDRIIMNNMTMWE 575

Db 570 GIKQLQARVLAVERYLGDQQLLGIGWCSGSKLICCTAVPWNASWSNKSIDRIIMNNMTMWE 629

Qy 576 GIKQLQARVLAVERYLGDQQLLGIGWCSGSKLICCTAVPWNASWSNKSIDRIIMNNMTMWE 675

Db 630 DREINNTYSLHSLIEESQNOEKNEQBELBLDKASIMNFDITKMLY 679

RESULT 9

env polypeptide - human immunodeficiency virus type 1 (isolate BR)
 N:Alternate names: coat polypeptide
 N:Contains: coat protein gp120; coat protein gp41
 C:Species: human immunodeficiency virus type 1, HIV-1
 A:Note: host Homo sapiens (man)
 C:Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 16-Feb-1997
 C:Accession: A31667
 R:Andand, R.; Thayer, R.; Srinivasan, A.; Nayyar, S.; Gardner, M.; Luciw, P.; Dandekar, S.
 Virology 168, 79-89, 1989
 A:Title: Biological and molecular characterization of human immunodeficiency virus (HIV-1)
 A:Reference number: A94389; PMID:89085613; PMID:2789516
 A:Accession: A31667
 A:Molecule type: DNA
 A:Residues: 1-852 <ANA>
 C:Superfamily: type E retrovirus env polypeptide
 C:Keywords: capsid protein; coat protein; polypeptide; transmembrane protein
 F:1-516/Product: coat protein gp120 #status predicted <CPI>
 F:517-852/Product: coat protein gp41 #status predicted <CPI>

Query Match 85.8%; Score 2875; DB 1; Length 852;
 Best Local Similarity 83.6%; Pred. No. 6, 7e-206;

Matches 541; Conservative 39; Mismatches 43; Indels 24; Gaps 5;

Qy 2 EKLWTVYVGVVPMKEATTLTFCASDAKAYDTEVHNWATACVPTDNPQEVLENYTE 61

Db 32 EKLWTVYVGVVPMKEATTLTFCASDAKAYDTEVHNWATACVPTDNPQEVLENYTE 91

Qy 62 HFNMMKNMVMQOMODIISLDQSLKPCVKLTPLCTVLANCKDVNATNTTNSSEGE--THER 121

Db 92 HFNMMKNMVMQOMODIISLDQSLKPCVKLTPLCTVLANCKDVNATNTTNSSEGE--THER 151

Qy 122 IKKCSFNITTSIRDEVOKEVALFYKLDVYKIDNN--NTSYRLISCDTSVITQACPKISFE 179

Db 152 IKKCSFNITTSIRDEVOKEVALFYKLDVYKIDNN--NTSYRLISCDTSVITQACPKISFE 211

Qy 160 PIPHYCAPAGFAILKCKNDKTFNGKXPCKNVSTXOCTGIRPVVSTOLLNGSLAEEV 239

Db 212 PIPHYCAPAGFAILKCKNDKTFNGKXPCKNVSTXOCTGIRPVVSTOLLNGSLAEEV 271

Qy 240 IRSDNFTNNAKTIIVQLKESVEINCTRPNNN-----GAGDIRQAHN 281

Db 272 IRSDNFTNNAKTIIVQLKESVEINCTRPNNN-----GAGDIRQAHN 331

Qy 282 ISRAKNDTLKQIVYKLRQEF--ENKTIYFNSSGGDEIVMHSFNGGEPFYCNSTOLF 341

Db 332 ISRAKNDTLKQIVYKLRQEF--ENKTIYFNSSGGDEIVMHSFNGGEPFYCNSTOLF 391

Qy 342 TWNNNTBESNTEGNT--ITLPCRKQIIMMOEVGKAMVAPPIRGOIRCSNITG 400

Db 392 TWNNNTBESNTEGNT--ITLPCRKQIIMMOEVGKAMVAPPIRGOIRCSNITG 450

Qy 401 DQGINE--NGTEIFRPGGDMRDNMRSELYKKYKVKIEPLGVAPTKRRVYQREK 458

Db 451 DQGINE--NGTEIFRPGGDMRDNMRSELYKKYKVKIEPLGVAPTKRRVYQREK 510

Qy 459 IGAVLGLGAGSTMGASMTLTVOARLLLSGIYQOQNNLLRAIEAOQRMQLTWG 518

Db 511 IGAVLGLGAGSTMGASMTLTVOARQLSGIYQOQNNLLRAIEAOQHLQLTW 570

Qy 519 QLOARVLAVERYLGDQQLLGIGWCSGSKLICCTAVPWNASWSNKSIDRIIMNNMTMWE 578

Db 571 QLOARVLAVERYLGDQQLLGIGWCSGSKLICCTAVPWNASWSNKSIDRIIMNNMTMWE 630

Qy 579 IDNTSEIYTLIEESQNOEKNEQBELBLDKASIMNFDITKMLY 625

Db 631 IDNTNLIYSLIEDSQIQOEKNEKELLELDKMASLMMWFNTNMLWY 677

RESULT 10

533985
env polyprotein - human immunodeficiency virus type 1
C/Species: human immunodeficiency virus type 1, HIV-1
C/Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 26-Aug-1999
C/Accession: S33985
R/Carlini, F.
Submitted to the EMBL Data Library, November 1991
A/Reference number: S33979
A/Accession: S33985
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-851 <CAR>
A/Cross-references: EMBL:Z11530, NID:960192, PIDN:CAA77628.1, PID:960199
C/Superfamily: type E retrovirus env polyprotein

Query Match 85.5%; Score 2866.5; DB 2; Length 851;
Best Local Similarity 83.9%; Pred. No. 2.9e-205;
Matches 543; Conservative 33; Mismatches 46; Indels 25; Gaps 5;

Qy 2 EKLMTVTVYGVVPWKEATITLFCASDAKAYDTEVHNWATHACVPTDPNPQEVLENT 61
Db 32 EKLMTVTVYGVVPWKEATITLFCASDAKAYDTEVHNWATHACVPTDPNPQEVLENT 91
Qy 62 HENMKNNWVQWQEDIIISLMDQSLKPCVKLTPLCYTLNCKDVNAINTNDSEG--TWER 119
Db 92 HENMKNNWVQWQEDIIISLMDQSLKPCVKLTPLCYTLNCKDVNAINTNDSEG--TWER 151
Qy 120 GEINCSFNITTSIRDEYQKEALFYKLDVYXIDNNNTSYRLISCDTSVITQACPKISFE 179
Db 152 GEINCSFNITTSIRDEYQKEALFYKLDVYXIDNNNTSYRLISCDTSVITQACPKISFE 211
Qy 180 PIPHYCAPAGFALILKCNKTFNGKXPKXNVSTXQCTHGIRPVVSTQLLNGSLAE 239
Db 212 PIPHYCAPAGFALILKCNKTFNGKXPKXNVSTXQCTHGIRPVVSTQLLNGSLAE 271
Qy 240 IRSNPFNNNAKTIIVOLKESVEINCTRPNN-----GAGDIRQAH 280
Db 272 IRSNPFNNNAKTIIVOLKESVEINCTRPNN-----GAGDIRQAH 331
Qy 281 NISPAKNNNDLTKQIVILKREOF-ENKTIIVFNHSSGDEIYVHSPNCGEFPFYCNSTQ 339
Db 332 NISPAKNNNDLTKQIVILKREOF-ENKTIIVFNHSSGDEIYVHSPNCGEFPFYCNSTQ 391
Qy 340 NSTWNNNTSGSNTEG--NTITLPCRIRKQIIMMOEGKAMYPAPIRGQIRCSSNITGLL 398
Db 392 NSTW--STKGSNNTEGSDITLPCRIRKQIIMMOEGKAMYPAPISGQIRCSSNITGLL 449
Qy 399 TRDGINENGTEIRFPGGDMRDWNRSELKYKVKLEPLGVAATKCRARVVOEKAVG 458
Db 450 TRDGSNNTEIRFPGGDMRDWNRSELKYKVKLEPLGVAATKCRARVVOEKAVG 509
Qy 459 IGAFLFLGAGSTGMAASMTLTVQARLLLSGIVOOQNNLLRAIEAQORMLQTLWG 518
Db 510 IGAFLFLGAGSTGMAASMTLTVQARLLLSGIVOOQNNLLRAIEAQORMLQTLWG 569
Qy 519 QLOARVLAVERYLDDQQLGIWGSGLICTTAVPWNASNSKSLDRIIMNMTEWERE 578
Db 570 QLOARVLAVERYLDDQQLGIWGSGLICTTAVPWNASNSKSLDRIIMNMTEWERE 629
Qy 579 IDNTSSEIYTLIESQNOEKNEOELLELDKMASLMMWFNTNMLWY 625
Db 630 IDNTSSEIYTLIESQNOEKNEOELLELDKMASLMMWFNTNMLWY 676

RESULT 11

VCLJ3M
env polyprotein precursor - human immunodeficiency virus type 1 (isolate WMJ1)
N/Contains: coat protein gp120; coat protein gp41

C/Species: human immunodeficiency virus type 1, HIV-1
C/Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 07-Nov-1997
C/Accession: A24774
R/Starchich, B.R.; Hahn, B.H.; Shaw, G.M.; McNeely, P.D.; Modrow, S.; Wolf, H.; Parks, E.
Cell 45; 637-648, 1986
A/Title: Identification and characterization of conserved and variable regions in the
A/Reference number: A24774; MUID:86218077; PMID:2423250
A/Accession: A24774
A/Molecule type: DNA
A/Residues: 1-856 <SPA>
A/Cross-references: GB:K03455, GB:M38432, NID:91906382
C/Genetics:
A/Gene: env
C/Superfamily: type E retrovirus env polyprotein
C/Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein
F.1-29/Domain: signal sequence #status predicted <SIG>
F.30-501/Product: coat protein gp120 #status predicted <GP1>
F.502-847/Product: coat protein gp41 #status predicted <GP2>
F.87,134,140,151,155,183,197,234,241,262,276,289,295,331,338,354,360,390,394,404,447,459

Query Match 85.4%; Score 2861.5; DB 1; Length 856;
Best Local Similarity 83.6%; Pred. No. 6.8e-205;
Matches 547; Conservative 29; Mismatches 47; Indels 31; Gaps 6;

Qy 1 VEKLMVTYVYGVVPWKEATITLFCASDAKAYDTEVHNWATHACVPTDPNPQEVLENT 60
Db 30 VEKLMVTYVYGVVPWKEATITLFCASDAKAYDTEVHNWATHACVPTDPNPQEVLENT 89
Qy 61 EHRPMKNNWVQWQEDIIISLMDQSLKPCVKLTPLCYTLNCKDVNAINTNDSEG 120
Db 90 EHRPMKNNWVQWQEDIIISLMDQSLKPCVKLTPLCYTLNCKDVNAINTNDSEG 147
Qy 121 EIKRCSFNITTSIRDEYQKEALFYKLDVYXIDNNNTSYRLISCDTSVITQACPK 175
Db 148 EIKRCSFNITTSIRDEYQKEALFYKLDVYXIDNNNTSYRLISCDTSVITQACPK 207
Qy 176 ISPEPIHYCAPAGFALILKCNKTFNGKXPKXNVSTXQCTHGIRPVVSTQLLNGSLAE 235
Db 208 ISPEPIHYCAPAGFALILKCNKTFNGKXPKXNVSTXQCTHGIRPVVSTQLLNGSLAE 267
Qy 236 BEVYIRSDNTNNAKTIIVOLKESVEINCTRPNN-----AGDIRQ 277
Db 268 BEVYIRSDNTNNAKTIIVOLKESVEINCTRPNN-----AGDIRQ 327
Qy 278 AHGNSAKNNNDLTKQIVILKREOF-ENKTIIVFNHSSGDEIYVHSPNCGEFPFYCNSTQ 337
Db 328 AHGNSAKNNNDLTKQIVILKREOF-ENKTIIVFNHSSGDEIYVHSPNCGEFPFYCNSTQ 387
Qy 338 LFNSTWN--NTEGSNNTE--GNTITLPCRIRKQIIMMOEGKAMYPAPIRGQIRCSSN 392
Db 388 LFNSTWNVTGISTEGNNNTENGSDITLPCRIRKQIIMMOEGKAMYPAPISGQIRCSSN 447
Qy 393 ITGILLTRDGINENGTEIRFPGGDMRDWNRSELKYKVKLEPLGVAATKCRARV 452
Db 448 ITGILLTRDGSNNSSREIRFPGGDMRDWNRSELKYKVKLEPLGVAATKCRARV 507
Qy 453 EKRAVIG-IGAVFLGFLGAGSTGMAASMTLTVQARLLLSGIVOOQNNLLRAIEAQORMLQ 511
Db 508 EKRAVIG-IGAVFLGFLGAGSTGMAASMTLTVQARLLLSGIVOOQNNLLRAIEAQORMLQ 567
Qy 512 LTVWGIQLOARVLAVERYLDDQQLGIWGSGLICTTAVPWNASNSKSLDRIIMNM 571
Db 568 LTVWGIQLOARVLAVERYLDDQQLGIWGSGLICTTAVPWNASNSKSLDRIIMNM 627
Qy 572 WMEWEREIDNTSSEIYTLIESQNOEKNEOELLELDKMASLMMWFNTNMLWY 625
Db 628 WMEWEREIDNTSSEIYTLIESQNOEKNEOELLELDKMASLMMWFNTNMLWY 681

RESULT 12

VCLJMN
env polyprotein precursor - human immunodeficiency virus type 1 (isolate MN)
N/Alternate names: coat polyprotein

C:Species: human immunodeficiency virus type 1, HIV-1
 A:Note: host Homo sapiens (man)
 C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Feb-1997
 C:Accession: A28922
 R:Gungo, C.; Guo, H.G.; Franchini, G.; Aldovini, A.; Collalti, E.; Farrell, K.; Wong-Sta
 virology 164, 531-536, 1988
 A>Title: Envelope sequences of two new United States HIV-1 isolates.
 A:Reference number: A28922; MUID:88219542; PMID:3369091
 A:Accession: A28922
 A:Molecule type: DNA
 A:Residues: 1-859 <GUR>
 C:Genetics:
 A:Gene: env
 C:Superfamily: type E retrovirus env polypeptide
 C:Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane prote
 F.1-29/Domain: signal sequence #status predicted <SIG>
 F.30-859/Product: env polypeptide #status predicted <EP>
 F:87,129,135,140,141,146,161,165,191,202,246,267,281,294,300,336,343,359,365,389,395,401

Query Match 85.3%; Score 2859.5; DB 1; Length 859;
 Best Local Similarity 83.1%; Pred. No. 9.7e-205;

Matches 545; Conservative 32; Mismatches 46; Indels 33; Gaps 7;

```

QY      2 EKLAVTVYGVVWKEATTTTFCASDAKAYDTEVHNVWATNACVPTDPNPOEVLNATE 61
DB      31 EKLAVTVYGVVWKEATTTTFCASDAKAYDTEVHNVWATNACVPTDPNPOEVLNATE 90
QY      62 HFNMMKNMVEOMEDIIISLDOSLPCVKLTPLCVTLNCDV-----NATNTNDSG 113
DB      91 NFNMMKNMVEOMEDIIISLDOSLPCVKLTPLCVTLNCDV-----NATNTNDSG 150
QY      114 EGTMERGEIKNCSFNITTSIRDEVOKEYALFYKLDVXIDNNNTSYRLISCDTSVITQAC 173
DB      151 EGTIGEMKNCSPNITTSIRDEVOKEYALFYKLDVXIDNNNTSYRLISCDTSVITQAC 210
QY      174 PKISPEPIPIHYCAPAGFALIKNDKTRNGKPCKNVSTQCTGIRPVSTOLLNGSL 233
DB      211 PKISPEPIPIHYCAPAGFALIKNDKTRNGKPCKNVSTQCTGIRPVSTOLLNGSL 270
QY      234 AEEVVRISDNFTNNAKTIIVOLKESVEINCTRPN-----NAGADI 275
DB      271 AEEVVRISDNFTNNAKTIIVOLKESVEINCTRPN-----NAGADI 330
QY      276 ROAHGNISRAKNDTLKQIVIKLREOFENKTIIVFNHSGGDEIYMHSGFNGCEFFYCN 335
DB      331 ROAHGNISRAKNDTLKQIVIKLREOFENKTIIVFNHSGGDEIYMHSGFNGCEFFYCN 390
QY      336 TOLFNSTNNNTES--TEGSSNT-EGNTTILPCRIKQIIMMOEYGVKAMTAPPIRQIRCS 391
DB      391 TOLFNSTNNNTES--TEGSSNT-EGNTTILPCRIKQIIMMOEYGVKAMTAPPIRQIRCS 450
QY      392 NITGLLTRDGIENG-----TEIFRPGGDMNDMSSELYKRVVYIEPLGVAFTCK 446
DB      451 NITGLLTRDGIENG-----TEIFRPGGDMNDMSSELYKRVVYIEPLGVAFTCK 510
QY      450 VQREKRAVIGAVFLGFLGAAGSTWGAASMTLTVOARLLSGIYQOONNLRAIEAORM 509
DB      511 VQREKRAVIGAVFLGFLGAAGSTWGAASMTLTVOARLLSGIYQOONNLRAIEAORM 569
QY      510 LQLTWVGIKQIOLQARVLAVERYLGDQQLIGWCSGKLICTTAVPWNASWSNKSIDR 565
DB      570 LQLTWVGIKQIOLQARVLAVERYLGDQQLIGWCSGKLICTTAVPWNASWSNKSIDR 629
QY      570 MTNMEMEESINDYNTSEIYTLIEESONOEKQOELELDKVASLMMNFDTIKWLMY 625
DB      630 MTNMEMEESINDYNTSEIYTLIEESONOEKQOELELDKVASLMMNFDTIKWLMY 685

```

RESULT 13

VCLJ3C

env polypeptide precursor - human immunodeficiency virus type 1 (isolate SC)

N:Alternate names: coat polypeptide

C:Species: human immunodeficiency virus type 1, HIV-1

A:Note: host Homo sapiens (man)
 C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Feb-1997
 C:Accession: B28922
 R:Gungo, C.; Guo, H.G.; Franchini, G.; Aldovini, A.; Collalti, E.; Farrell, K.; Wong-Sta
 virology 164, 531-536, 1988
 A>Title: Envelope sequences of two new United States HIV-1 isolates.
 A:Reference number: A28922; MUID:88219542; PMID:3369091
 A:Accession: B28922
 A:Molecule type: DNA
 A:Residues: 1-861 <GUR>
 C:Genetics:
 A:Gene: env
 C:Superfamily: type E retrovirus env polypeptide
 C:Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane prote
 F.1-29/Domain: signal sequence #status predicted <SIG>
 F.30-861/Product: env polypeptide #status predicted <EP>
 F:87,129,135,140,143,159,163,187,198,224,241,262,276,295,301,302,333,340,356,362,386,396

Query Match 85.1%; Score 2853; DB 1; Length 861;
 Best Local Similarity 82.9%; Pred. No. 3e-204;

Matches 547; Conservative 32; Mismatches 41; Indels 40; Gaps 9;

```

QY      2 EKLAVTVYGVVWKEATTTTFCASDAKAYDTEVHNVWATNACVPTDPNPOEVLNATE 61
DB      31 EKLAVTVYGVVWKEATTTTFCASDAKAYDTEVHNVWATNACVPTDPNPOEVLNATE 90
QY      62 HFNMMKNMVEOMEDIIISLDOSLPCVKLTPLCVTLNCDV-----NATNTNDSG 115
DB      91 NFNMMKNMVEOMEDIIISLDOSLPCVKLTPLCVTLNCDV-----NATNTNDSG 150
QY      116 EGTMERGEIKNCSFNITTSIRDEVOKEYALFYKLDVXIDNNNTSYRLISCDTSVITQAC 175
DB      151 EGTIGEMKNCSPNITTSIRDEVOKEYALFYKLDVXIDNNNTSYRLISCDTSVITQAC 208
QY      176 ISPEPIPIHYCAPAGFALIKNDKTRNGKPCKNVSTQCTGIRPVSTOLLNGSLAE 235
DB      209 ISPEPIPIHYCAPAGFALIKNDKTRNGKPCKNVSTQCTGIRPVSTOLLNGSLAE 267
QY      236 EEEVVRISDNFTNNAKTIIVOLKESVEINCTRPNNGA-----GDI 275
DB      268 EEEVVRISDNFTNNAKTIIVOLKESVEINCTRPNNGA-----GDI 327
QY      276 ROAHGNISRAKNDTLKQIVIKLREOFENKTIIVFNHSGGDEIYMHSGFNGCEFFYCN 335
DB      338 ROAHGNISRAKNDTLKQIVIKLREOFENKTIIVFNHSGGDEIYMHSGFNGCEFFYCN 387
QY      336 TOLFNSTNNNTES--NTEGN-TTILPCRIKQIIMMOEYGVKAMTAPPIRQIRCS 391
DB      388 TOLFNSTNNNTES--NTEGN-TTILPCRIKQIIMMOEYGVKAMTAPPIRQIRCS 446
QY      392 NITGLLTRDGIENG-----TEIFRPGGDMNDMSSELYKRVVYIEPLGVAFTCK 446
DB      447 NITGLLTRDGIENG-----TEIFRPGGDMNDMSSELYKRVVYIEPLGVAFTCK 506
QY      447 RRVQREKRAVIGAVFLGFLGAAGSTWGAASMTLTVOARLLSGIYQOONNLRAIEA 505
DB      507 RRVQREKRAVIGAVFLGFLGAAGSTWGAASMTLTVOARLLSGIYQOONNLRAIEA 566
QY      506 QOQMLQLTWVGIKQIOLQARVLAVERYLGDQQLIGWCSGKLICTTAVPWNASWSNKSIDR 565
DB      567 QOQMLQLTWVGIKQIOLQARVLAVERYLGDQQLIGWCSGKLICTTAVPWNASWSNKSIDR 626
QY      566 IMNNNTWMEEREIDNTYNTSEIYTLIEESONOEKQOELELDKVASLMMNFDTIKWLMY 625
DB      627 IMNNNTWMEEREIDNTYNTSEIYTLIEESONOEKQOELELDKVASLMMNFDTIKWLMY 686

```

RESULT 14

VCLJ2

env polypeptide precursor - human immunodeficiency virus type 1 (isolate ARV-2)

N:Alternate names: coat polypeptide

A:Note: host Homo sapiens (man)

C>Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 16-Jul-1999
 C/Accession: A03976
 C/Species: human immunodeficiency virus type 1, HIV-1
 C/Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 16-Jul-1999
 C/Accession: C25523
 R/Debar, S.M.; Kalyanaraman, V.S.; Casey, J.M.; Strinivasan, A.; Andersen, P.R.; Devare, S.
 Proc. Natl. Acad. Sci. U.S.A. 83, 8380-8384, 1986
 A/Title: Molecular cloning and primary nucleotide sequence analysis of a distinct human
 A/Reference number: A94136; MUID:87041461; PMID:3490666
 A/Accession: C25523
 A/Molecule type: DNA
 A/Residues: 1-855 <SN>
 A/Cross-references: GB:K02007; NID:g328658; PIDN:AA59882.1; PID:g328666
 C/Genetics:
 A/Gene: env
 C/Superfamily: type B retrovirus env polyprotein
 C/Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprot
 F:1-30/Domain: signal sequence #status predicted <SIG>
 F:31-509/Product: exterior membrane glycoprotein #status predicted <EXT>
 F:510-855/Product: transmembrane glycoprotein #status predicted <TM>
 F:87,129,140,158,184,190,244,265,292,298,304,334,341,358,364,388,394,400,408,445,458
 F:610,624,636,815/Binding site: carbohydrate (Aam) (covalent) #status predicted

Query Match 84.6%; Score 2836; DB 1; Length 855;
 Best Local Similarity 83.2%; Pred. No. 5.4e-203;

Matches 543; Conservative 30; Mismatches 48; Indels 32; Gaps 9;

Qy 2 EKLWVTVYGVVWKEATTTTLCASDAKAYDTEVHNWATHACVPTDNPQEVLENTHEF 61
 Db 31 EKLWVTVYGVVWKEATTTTLCASDAKAYDTEVHNWATHACVPTDNPQEVLENTHEF 90
 Qy 62 HNNWNNWVQWQEDDIIISLMQSLKPCVKLTPLCYTLNCKVNAVNTNTDSEGMT- 120
 Db 91 HNNWNNWVQWQEDDIIISLMQSLKPCVKLTPLCYTLNCKVNAVNTNTDSEGMT- 150
 Qy 121 EIKKCSFNITTSIRDEVOKEVALFYKLDVYXIDN-----NNTSYRLISCDTSVITQCPK 175
 Db 151 EIKKCSFNITTSIRDEVOKEVALFYKLDVYXIDN-----NNTSYRLISCDTSVITQCPK 210
 Qy 176 ISFEPPIHYCAPAGFAILKCNKDTFNKGPCKNVSTXQCTHGIRPVVSTOLLNGSLAE 235
 Db 211 VSFEPPIHYCTPAGFAILKCNKDTFNKGPCKNVSTXQCTHGIRPVVSTOLLNGSLAE 270
 Qy 236 EEVYIRSDNFTNNAKTIIVQLKESVEINCTRPNNNG-----AGDIRQ 277
 Db 271 EEVYIRSDNFTNNAKTIIVQLKESVEINCTRPNNNG-----AGDIRQ 330
 Qy 278 AHGNSIRAKNDTLKOIYIKLREDF-ENKTIIVFNHSSGDEIYVHSHFNCGEFFYC 336
 Db 331 AHGNSIRAKNDTLKOIYIKLREDF-ENKTIIVFNHSSGDEIYVHSHFNCGEFFYC 390
 Qy 337 QLFNSTVN-NNTGSSNNTGEG-TTLPCRIOIINMVEGKAMVAPPPIRGQIRCSNIT 394
 Db 391 QLFNSTVN-NNTGSSNNTGEG-TTLPCRIOIINMVEGKAMVAPPPIRGQIRCSNIT 447
 Qy 395 GLLTTRDGIN-ENGTEIFRPGGDMRDNRSELYKRVVYKIEPLGVAFTCKRVORE 453
 Db 448 GLLTTRDGIN-ENGTEIFRPGGDMRDNRSELYKRVVYKIEPLGVAFTCKRVORE 507
 Qy 454 KRAVGI-IGAVPLGFGAAGTGAASMTLTVOARLLSGIYQOQNNLLRAIEAOQHILQ 512
 Db 508 KRAVGIIGAVPLGFGAAGTGAASMTLTVOARLLSGIYQOQNNLLRAIEAOQHILQ 567
 Qy 513 TVWGIGKQIQAARVLAVERYLQDQLLGWCGSKLICCTAVPWNASWSNKSIDRI 572
 Db 568 TVWGIGKQIQAARVLAVERYLQDQLLGWCGSKLICCTAVPWNASWSNKSIDRI 627
 Qy 573 MEMEREIDNYSSELYTLIEESQNOQEKQOGLLELDKASIMNFDTIKMLY 625
 Db 628 MEMEREIDNYSSELYTLIEESQNOQEKQOGLLELDKASIMNFDTIKMLY 680

RESULT 15
 VCLJH4
 env polyprotein - human immunodeficiency virus type 1 (isolate CDC-451)
 N/Alternate names: coat polyprotein

N/Contains: coat protein gp120; coat protein gp41
 C/Species: human immunodeficiency virus type 1, HIV-1
 C/Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 16-Jul-1999
 C/Accession: C25523
 R/Debar, S.M.; Kalyanaraman, V.S.; Casey, J.M.; Strinivasan, A.; Andersen, P.R.; Devare, S.
 Proc. Natl. Acad. Sci. U.S.A. 83, 8380-8384, 1986
 A/Title: Molecular cloning and primary nucleotide sequence analysis of a distinct human
 A/Reference number: A94136; MUID:87041461; PMID:3490666
 A/Accession: C25523
 A/Molecule type: DNA
 A/Residues: 1-868 <DES>
 A/Cross-references: GB:M13137; NID:g326460; PIDN:AAA4311.1; PID:g326467
 C/Genetics:
 A/Gene: env
 C/Superfamily: type B retrovirus env polyprotein
 C/Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane protei
 F:1-521/Product: coat protein gp120 #status predicted <GP1>
 F:522-868/Product: coat protein gp41 #status predicted <GP2>
 F:89,138,139,142,166,195,198,208,245,252,273,287,300,312,342,349,365,371,395,405,409,459,

Query Match 83.4%; Score 2793.5; DB 1; Length 868;
 Best Local Similarity 80.1%; Pred. No. 8.1e-200;

Matches 528; Conservative 36; Mismatches 58; Indels 37; Gaps 5;

Qy 4 LWTVYGVVWKEATTTTLCASDAKAYDTEVHNWATHACVPTDNPQEVLENTHEF 63
 Db 35 LWTVYGVVWKEATTTTLCASDAKAYDTEVHNWATHACVPTDNPQEVLENTHEF 94
 Qy 64 HNNWNNWVQWQEDDIIISLMQSLKPCVKLTPLCYTLNCKVNAVNTNTDSEGMT- 117
 Db 95 HNNWNNWVQWQEDDIIISLMQSLKPCVKLTPLCYTLNCKVNAVNTNTDSEGMT- 154
 Qy 118 -ERGEKNGSFNITTSIRDEVOKEVALFYKLDVYXIDN-----NNTSYRLISCDTSVITQ 171
 Db 155 -ERGEKNGSFNITTSIRDEVOKEVALFYKLDVYXIDN-----NNTSYRLISCDTSVITQ 214
 Qy 172 ACPKISFEPPIHYCAPAGFAILKCNKDTFNKGPCKNVSTXQCTHGIRPVVSTOLLNG 231
 Db 215 ACPKISFEPPIHYCTPAGFAILKCNKDTFNKGPCKNVSTXQCTHGIRPVVSTOLLNG 274
 Qy 232 SLAEEVYIRSDNFTNNAKTIIVQLKESVEINCTRPNNNG-----AG 273
 Db 275 SLAEEVYIRSDNFTNNAKTIIVQLKESVEINCTRPNNNG-----AG 334
 Qy 274 DIRAHGNISRAKNDTLKOIYIKLREDF-ENKTIIVFNHSSGDEIYVHSHFNCGEFFYC 333
 Db 335 DIRAHGNISRAKNDTLKOIYIKLREDF-ENKTIIVFNHSSGDEIYVHSHFNCGEFFYC 394
 Qy 334 NSTQLFNSTVNNNTGEG-----NNTGNTITLPCRIOIINMVEGKAMVAPPPIRGQ 387
 Db 395 NSTQLFNSTVNNNTGEG-----NNTGNTITLPCRIOIINMVEGKAMVAPPPIRGQ 454
 Qy 388 RCGSNTIGLLTTRDGIN-ENGTEIFRPGGDMRDNRSELYKRVVYKIEPLGVAFTCKR 447
 Db 455 RCGSNTIGLLTTRDGIN-ENGTEIFRPGGDMRDNRSELYKRVVYKIEPLGVAFTCKR 514
 Qy 448 RVVOREKRAVG-IGAVPLGFGAAGTGAASMTLTVOARLLSGIYQOQNNLLRAIEAO 506
 Db 515 RVVOREKRAVG-IGAVPLGFGAAGTGAASMTLTVOARLLSGIYQOQNNLLRAIEAO 574
 Qy 507 QRMQLTVWGIGKQIQAARVLAVERYLQDQLLGWCGSKLICCTAVPWNASWSNKSIDRI 566
 Db 575 QRMQLTVWGIGKQIQAARVLAVERYLQDQLLGWCGSKLICCTAVPWNASWSNKSIDRI 634
 Qy 567 WNNNTWMEEREIDNYSSELYTLIEESQNOQEKQOGLLELDKASIMNFDTIKMLY 625
 Db 635 WNNNTWMEEREIDNYSSELYTLIEESQNOQEKQOGLLELDKASIMNFDTIKMLY 693

Search completed: December 12, 2003, 12:33:34
 Job time : 15.5306 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: December 12, 2003, 12:26:14 ; Search time 32.1467 Seconds
(without alignments)
5017.086 Million cell updates/sec

Title: US-10-032-162-17
Perfect score: 3351
Sequence: 1 VEKLWVTVYGVVPMKREAT.....ELDKWASIMNWFDTKMLWY 625

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
1 number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3293	98.3	847	15	Q75760 human immun
2	2993.5	89.3	848	15	Q74999 human immun
3	2944.5	87.9	852	15	Q92761 human immun
4	2944.5	87.9	854	15	Q78225 human immun
5	2939.5	87.7	846	15	Q9PXW7 human immun
6	2938.5	87.7	855	15	003805 human immun
7	2936.5	87.6	854	15	040222 human immun
8	2935.5	87.6	838	15	003806 human immun
9	2931	87.5	849	15	Q77368 human immun
10	2929	87.4	854	15	Q92762 human immun
11	2928	87.3	853	15	Q77989 human immun
12	2926	87.3	853	15	003811 human immun
13	2924	87.3	851	15	056110 human immun
14	2917	87.0	847	15	Q8AR21 human immun
15	2911	86.9	849	15	Q8Q851 human immun
16	2906.5	86.7	860	15	Q9YP50 human immun

17	2903.5	86.6	868	15	Q9WJW5 human immun
18	2900.5	86.6	852	15	Q41883 human immun
19	2900	86.5	861	15	Q9WJW5 human immun
20	2899	86.5	855	15	Q8U644 human immun
21	2899	86.5	859	15	Q8Q850 human immun
22	2898	86.5	873	15	Q69994 human immun
23	2896	86.4	843	15	Q70150 human immun
24	2896	86.4	851	15	Q56562 human immun
25	2895	86.4	752	15	Q70604 human immun
26	2895	86.4	752	15	Q70605 human immun
27	2895	86.4	861	15	Q9E527 human immun
28	2894.5	86.4	856	15	Q71270 human immun
29	2894	86.4	861	15	Q9E523 human immun
30	2892	86.3	853	15	Q56108 human immun
31	2892	86.3	863	15	Q42031 human immun
32	2889	86.2	854	15	Q72502 human immun
33	2889	86.2	863	15	Q9WJW8 human immun
34	2887.5	86.2	850	15	Q74812 human immun
35	2887.5	86.2	856	15	Q72993 human immun
36	2886	86.1	854	15	Q85582 human immun
37	2885.5	86.1	840	15	Q8Q2X4 human immun
38	2885.5	86.1	850	15	Q56113 human immun
39	2884.5	86.1	860	15	Q9YF43 human immun
40	2884.5	86.1	860	15	Q9DH01 human immun
41	2884	86.1	851	15	Q8Q852 human immun
42	2883.5	86.0	843	15	Q70008 human immun
43	2881	86.0	856	15	Q92877 human immun
44	2881	86.0	863	15	Q9WJW4 human immun
45	2880.5	86.0	843	15	Q9QK48 human immun

ALIGNMENTS

RESULT 1
ID Q75760 PRELIMINARY; PRT; 847 AA.
AC Q75760;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Envelope glycoprotein.
EN
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_taxid=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JREF;
RX MEDLINE=91043044; PubMed=2172833;
RA O'Brien W.A., Koyanagi Y., Namazie A., Zhao J.Q., Diagne A., Idler K., Koyanagi Y., Miles S., Mitsuyasu R.T., Merrill J.E., Vinters H.V., Chen I.S.;
RT "Dual infection of the central nervous system by AIDS viruses with distinct cellular tropisms.";
RL Science 236:819-822(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=JREF;
RX MEDLINE=91043044; PubMed=2172833;
RA O'Brien W.A., Koyanagi Y., Namazie A., Zhao J.Q., Diagne A., Idler K., Zack J.A., Chen I.S.;
RT "HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120 outside the CD4-binding domain.";
RL Nature 348:69-73(1990).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=JREF;
RX MEDLINE=92092169; PubMed=1684385;
RA Pang S., Vinters H.V., Akashi T., O'Brien W.A., Chen I.S.;
RT "HIV-1 env sequence variation in brain tissue of patients with AIDS-related neurologic disease.";
RL J. Acquir. Immune Defic. Syndr. 4:1082-1092(1991).
RN [4]

RP SEQUENCE FROM N.A.
 RC STRAIN=JREFL;
 RA Pang S., Vintere H.V., Akashi T., O'Brien W.A., Chen I.S.,
 RA Koyanagi Y., Namazie A., Zhao J., Diagne A., Idler K.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U63632; AAB05604.1;
 DR InterPro; IPR000328; Env_GP41.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 KM AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
 SQ SEQUENCE 847 AA; 96160 MW; 022D5F24E04FB29P CRC64;

Query Match 98.3%; Score 3293; DB 15; Length 847;
 Best Local Similarity 96.1%; Pred. No. 9.1e-261;
 Matches 618; Conservative 0; Mismatches 7; Indels 18; Gaps 1;

QY 1 VEKLMVTVVYGVVWKEATTTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENVT 60
 VEKLMVTVVYGVVWKEATTTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENVT 89
 QY 61 EHFMMKNNWEOQEDIIISMDOSLKPCKVLTPLCYTLNCKVNAVTNTTDSGTMERG 120
 EHFMMKNNWEOQEDIIISMDOSLKPCKVLTPLCYTLNCKVNAVTNTTDSGTMERG 149
 QY 90 EHFMMKNNWEOQEDIIISMDOSLKPCKVLTPLCYTLNCKVNAVTNTTDSGTMERG 149
 EHKCSFNITTSIRDEQKAYALFYKLDVYKIDNNNTSYRLISGDTSVITQACPKISFE 180
 EHKCSFNITTSIRDEQKAYALFYKLDVYKIDNNNTSYRLISGDTSVITQACPKISFE 209
 QY 181 IPIHYCAPAGFALIKCNDKTFNGKPKCKNVSTXOCTHGIRPVVSTOLLNGSLAEDEV 240
 IPIHYCAPAGFALIKCNDKTFNGKPKCKNVSTXOCTHGIRPVVSTOLLNGSLAEDEV 269
 QY 210 IPIHYCAPAGFALIKCNDKTFNGKPKCKNVSTXOCTHGIRPVVSTOLLNGSLAEDEV 269
 IPIHYCAPAGFALIKCNDKTFNGKPKCKNVSTXOCTHGIRPVVSTOLLNGSLAEDEV 298
 QY 241 RSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIOAHCN 282
 RSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIOAHCN 329
 QY 270 RSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIOAHCN 329
 RSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIOAHCN 358
 QY 283 SRAKMNITLKOIVKLRQEFNKTIIVFNHSGGDEIIVHSGGDEFFYCNSTOLFNS 342
 SRAKMNITLKOIVKLRQEFNKTIIVFNHSGGDEIIVHSGGDEFFYCNSTOLFNS 389
 QY 330 SRAKMNITLKOIVKLRQEFNKTIIVFNHSGGDEIIVHSGGDEFFYCNSTOLFNS 389
 SRAKMNITLKOIVKLRQEFNKTIIVFNHSGGDEIIVHSGGDEFFYCNSTOLFNS 418
 QY 343 WNNTEGSNNTEGNTITLPCRIKQIINMOEVGKAMAPPIRGQIRGSSNITGLLTRDG 402
 WNNTEGSNNTEGNTITLPCRIKQIINMOEVGKAMAPPIRGQIRGSSNITGLLTRDG 449
 QY 390 WNNTEGSNNTEGNTITLPCRIKQIINMOEVGKAMAPPIRGQIRGSSNITGLLTRDG 449
 WNNTEGSNNTEGNTITLPCRIKQIINMOEVGKAMAPPIRGQIRGSSNITGLLTRDG 478
 QY 403 GINENGTEIFRPGGDMRDMNRSELYKKYKIVKIEPLGVAPTKARRRVQREKRAVGIG 462
 GINENGTEIFRPGGDMRDMNRSELYKKYKIVKIEPLGVAPTKARRRVQREKRAVGIG 509
 QY 450 GINENGTEIFRPGGDMRDMNRSELYKKYKIVKIEPLGVAPTKARRRVQREKRAVGIG 509
 GINENGTEIFRPGGDMRDMNRSELYKKYKIVKIEPLGVAPTKARRRVQREKRAVGIG 538
 QY 463 FLGFLGAAGSTMGASMTLTVQARLLLSGIVQOONNLRAIEAOORMLQTLVWGIKOLA 522
 FLGFLGAAGSTMGASMTLTVQARLLLSGIVQOONNLRAIEAOORMLQTLVWGIKOLA 569
 QY 510 FLGFLGAAGSTMGASMTLTVQARLLLSGIVQOONNLRAIEAOORMLQTLVWGIKOLA 569
 FLGFLGAAGSTMGASMTLTVQARLLLSGIVQOONNLRAIEAOORMLQTLVWGIKOLA 598
 QY 523 RVLAVERYLDQOLIGWCGSKLICCTAVPMNWSWSKSLDRIMNNNTWMEWEEREIDNY 582
 RVLAVERYLDQOLIGWCGSKLICCTAVPMNWSWSKSLDRIMNNNTWMEWEEREIDNY 629
 QY 570 RVLAVERYLDQOLIGWCGSKLICCTAVPMNWSWSKSLDRIMNNNTWMEWEEREIDNY 629
 RVLAVERYLDQOLIGWCGSKLICCTAVPMNWSWSKSLDRIMNNNTWMEWEEREIDNY 658
 QY 583 TSEIYTLIESQNOEKNEQELBLDKVASIMNFDITKMLWY 625
 TSEIYTLIESQNOEKNEQELBLDKVASIMNFDITKMLWY 672
 QY 630 TSEIYTLIESQNOEKNEQELBLDKVASIMNFDITKMLWY 672
 TSEIYTLIESQNOEKNEQELBLDKVASIMNFDITKMLWY 701

RESULT 2

QY 074999 PRELIMINARY; PRT; 848 AA.
 AC Q74999;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
 DE Envelope glycoprotein gp160.
 GN ENV.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OX NCBI_TaxId=11676;
 RN [1].
 RP SEQUENCE FROM N.A.
 RC STRAIN=JR-CSF;
 RX MEDLINE=97063128; PubMed=8906996;
 RA Klasse P.J., Boyd M.T., Weiss R.A., Schulz T.F.;
 RT "Mutations in the vpu, env, and nef genes of a syncytium-inducing
 RT variant of HIV type 1 JR-CSF that infects a range of T cell lines";
 RL AIDS Res. Hum. Retroviruses 12:347-350(1996).
 DR EMBL; U45960; AAB60591.1;
 DR InterPro; IPR000328; Env_GP41.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 KM AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
 SQ SEQUENCE 848 AA; 96433 MW; B240019C8737BCB3 CRC64;

Query Match 89.3%; Score 2993.5; DB 15; Length 848;
 Best Local Similarity 87.8%; Pred. No. 3.1e-236;
 Matches 566; Conservative 26; Mismatches 32; Indels 21; Gaps 4;

QY 1 VEKLMVTVVYGVVWKEATTTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENVT 60
 VEKLMVTVVYGVVWKEATTTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENVT 89
 QY 30 VEKLMVTVVYGVVWKEATTTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENVT 89
 VEKLMVTVVYGVVWKEATTTTLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLENVT 118
 QY 61 EHFMMKNNWEOQEDIIISMDOSLKPCKVLTPLCYTLNCKVNAVTNTTDSGTMERG 120
 EHFMMKNNWEOQEDIIISMDOSLKPCKVLTPLCYTLNCKVNAVTNTTDSGTMERG 149
 QY 90 EHFMMKNNWEOQEDIIISMDOSLKPCKVLTPLCYTLNCKVNAVTNTTDSGTMERG 149
 EHKCSFNITTSIRDEQKAYALFYKLDVYKIDNNNTSYRLISGDTSVITQACPKISFE 179
 EHKCSFNITTSIRDEQKAYALFYKLDVYKIDNNNTSYRLISGDTSVITQACPKISFE 209
 QY 150 EHKCSFNITTSIRDEQKAYALFYKLDVYKIDNNNTSYRLISGDTSVITQACPKISFE 209
 EHKCSFNITTSIRDEQKAYALFYKLDVYKIDNNNTSYRLISGDTSVITQACPKISFE 238
 QY 180 PIPHYCAPAGFALIKCNDKTFNGKPKCKNVSTXOCTHGIRPVVSTOLLNGSLAEDEV 239
 PIPHYCAPAGFALIKCNDKTFNGKPKCKNVSTXOCTHGIRPVVSTOLLNGSLAEDEV 269
 QY 210 PIPHYCAPAGFALIKCNDKTFNGKPKCKNVSTXOCTHGIRPVVSTOLLNGSLAEDEV 269
 PIPHYCAPAGFALIKCNDKTFNGKPKCKNVSTXOCTHGIRPVVSTOLLNGSLAEDEV 298
 QY 240 IRSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----GAGDIOAHCN 281
 IRSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----GAGDIOAHCN 329
 QY 270 IRSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----GAGDIOAHCN 329
 IRSDNFTNNAKTIIVOLKESVEINCTRPNNNG-----GAGDIOAHCN 358
 QY 282 ISRAKMNITLKOIVKLRQEFNKTIIVFNHSGGDEIIVHSGGDEFFYCNSTOLFNS 341
 ISRAKMNITLKOIVKLRQEFNKTIIVFNHSGGDEIIVHSGGDEFFYCNSTOLFNS 389
 QY 330 ISRAKMNITLKOIVKLRQEFNKTIIVFNHSGGDEIIVHSGGDEFFYCNSTOLFNS 389
 ISRAKMNITLKOIVKLRQEFNKTIIVFNHSGGDEIIVHSGGDEFFYCNSTOLFNS 418
 QY 342 WNNTEGSNNTEG-NTITLPCRIKQIINMOEVGKAMAPPIRGQIRGSSNITGLLTR 400
 WNNTEGSNNTEG-NTITLPCRIKQIINMOEVGKAMAPPIRGQIRGSSNITGLLTR 448
 QY 390 TW-NDEKSGTGEGYDITLPCRIKQIINMOEVGKAMAPPIRGQIRGSSNITGLLTR 448
 TW-NDEKSGTGEGYDITLPCRIKQIINMOEVGKAMAPPIRGQIRGSSNITGLLTR 477
 QY 401 DGINENGTEIFRPGGDMRDMNRSELYKKYKIVKIEPLGVAPTKARRRVQREKRAVGIG 460
 DGINENGTEIFRPGGDMRDMNRSELYKKYKIVKIEPLGVAPTKARRRVQREKRAVGIG 507
 QY 449 DGINENGTEIFRPGGDMRDMNRSELYKKYKIVKIEPLGVAPTKARRRVQREKRAVGIG 508
 DGINENGTEIFRPGGDMRDMNRSELYKKYKIVKIEPLGVAPTKARRRVQREKRAVGIG 537
 QY 461 AVFLGFLGAAGSTMGASMTLTVQARLLLSGIVQOONNLRAIEAOORMLQTLVWGIKOLA 520
 AVFLGFLGAAGSTMGASMTLTVQARLLLSGIVQOONNLRAIEAOORMLQTLVWGIKOLA 567
 QY 509 AVFLGFLGAAGSTMGASMTLTVQARLLLSGIVQOONNLRAIEAOORMLQTLVWGIKOLA 568
 AVFLGFLGAAGSTMGASMTLTVQARLLLSGIVQOONNLRAIEAOORMLQTLVWGIKOLA 597
 QY 521 QARVLAVERYLDQOLIGWCGSKLICCTAVPMNWSWSKSLDRIMNNNTWMEWEEREID 580
 QARVLAVERYLDQOLIGWCGSKLICCTAVPMNWSWSKSLDRIMNNNTWMEWEEREID 627
 QY 569 QARVLAVERYLDQOLIGWCGSKLICCTAVPMNWSWSKSLDRIMNNNTWMEWEEREIE 628
 QARVLAVERYLDQOLIGWCGSKLICCTAVPMNWSWSKSLDRIMNNNTWMEWEEREIE 657
 QY 581 NYSIYTLIESQNOEKNEQELBLDKVASIMNFDITKMLWY 625
 NYSIYTLIESQNOEKNEQELBLDKVASIMNFDITKMLWY 672
 QY 629 NYSIYTLIESQNOEKNEQELBLDKVASIMNFDITKMLWY 673
 NYSIYTLIESQNOEKNEQELBLDKVASIMNFDITKMLWY 702

RESULT 3

092761 PRELIMINARY; PRT; 852 AA.
 ID 092761
 AC 092761
 DT 01-NOV-1998 (TRENBLrel. 08, Created)
 DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Envelope glycoprotein.
 GN ENV.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SFMS2.9;
 RX MEDLINE=98178716; PubMed=9519894;
 RA McCutchan F.B., Sanders-Buell E., Salminen M.O., Carr J.K.,
 RA Shepard W.H.;
 "Diversity of the human immunodeficiency virus type 1 envelope
 glycoprotein in San Francisco Men's Health Study participants";
 AIDS Res. Hum. Retroviruses 14:329-337(1998).
 DR EMBL, AF025750; AAC40588.1; -;
 DR InterPro; IPR000328; Env_GP41.
 DR InterPro; IPR000173; GAP_dhdrogenase.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 DR PROSITE; PS00071; GAPDH; 1.
 DR AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
 SQ SEQUENCE 852 AA; 96545 MW; 3C7780DB0611B617 CRC64;

Query Match 87.9%; Score 2944.5; DB 15; Length 852;

Best Local Similarity 86.1%; Pred. No. 3.2e-232; Indels 25; Gaps 6;

Matches 557; Conservative 25; Mismatches 40; Indels 25; Gaps 6;

QY 3 KLMVTVYVGVVWKEATITLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNENYTEH 62
 DB 32 QLMVTVYVGVVWKEATITLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNENYTEH 91
 QY 63 FNMWKNMVEQMEDIIISLMDOSLKPCKVLTPLCVTLNCKDV-NATNTTNDSEGMERGE 121
 DB 92 FNMWKNMVEQMEDIIISLMDOSLKPCKVLTPLCVTLNCKDV-NATNTTNDSEGMERGE 151
 QY 122 IKNSFNITTSIRDEVOKEVALFYKLDVYXID--NNNTSYRLISCDTSVITQACPKISFE 179
 DB 152 IKNSFNITTSIRDEVOKEVALFYKLDVYXID--NNNTSYRLISCDTSVITQACPKISFE 211
 QY 180 PIPHYCAPAGFALIKCNDKTFNGKCKPCKNVSTXOCTHGIRPVVSTOLLNGSLAESEV 239
 DB 212 PIPHYCAPAGFALIKCNDKTFNGKCKPCKNVSTXOCTHGIRPVVSTOLLNGSLAESEV 271
 QY 240 IRSDFNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRQAHCN 281
 DB 272 IRSDFNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRQAHCN 331
 QY 282 ISRAKNDTLKQIVIKAREOFENKTIIVNHSOGDEIYVHSPNCGGEFFYCNSTQULNS 341
 DB 332 ISRAKNDTLKQIVIKAREOFENKTIIVNHSOGDEIYVHSPNCGGEFFYCNSTQULNS 391
 QY 342 TNNNTTSGNNTTEGN--TITLPCRIRKQIINMVEGKAMVAPPIRGQIRCSNITGLLT 399
 DB 392 TM-NDTESNNTRENDTITITPCRIRKQIINMVEGKAMVAPPIRGQIRCSNITGLLT 450
 QY 400 RDGGINENG-TEIFPPGGGDMRDNMRSELVYKVKVIEPLGVAPTKCRRVVQREKRAVG 458
 DB 451 RDGGINENG-TEIFPPGGGDMRDNMRSELVYKVKVIEPLGVAPTKCRRVVQREKRAVG 510
 QY 459 IGAVFLGLGAAGSTMGASMTLTVQARLLLSGIVQOQNNLLRAIEAOQMLQLTWGIX 518
 DB 511 IGAVFLGLGAAGSTMGASMTLTVQARLLLSGIVQOQNNLLRAIEAOQMLQLTWGIX 570
 QY 519 QLOARVLAVERYLDQDLGIGWCGSKLICCTAVPWNASWSNKSIDRIIMNNTWMEWRE 578

RESULT 4

DB 571 QLOARVLAVERYLDQDLGIGWCGSKLICCTAVPWNASWSNKSIDRIIMNNTWMEWRE 630
 QY 579 IDNTTSELYTILIESNOQEKNEBELDPMASLWAFDITTKLWY 625
 DB 631 IDNTTSELYTILIESNOQEKNEBELDPMASLWAFDITTKLWY 677
 ID 078225 PRELIMINARY; PRT; 854 AA.
 AC 078225;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
 DE Env protein.
 GN ENV.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Ba-1;
 RX MEDLINE=91289160; PubMed=1905842;
 RA Hwang S.S., Boyle T.J., Lyeily H.K., Cullen B.R.;
 RT "Identification of the envelope V3 loop as the primary determinant of
 RT cell tropism in HIV-1";
 RL Science 253:71-74(1991).
 DR EMBL; M63929; AAA75116.1; -;
 DR InterPro; IPR000328; Env_GP41.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 DR AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
 SQ SEQUENCE 854 AA; 97186 MW; 005FC8B1B5C8DF08 CRC64;

Query Match 87.9%; Score 2944.5; DB 15; Length 854;

Best Local Similarity 85.5%; Pred. No. 3.2e-232; Indels 25; Gaps 3;

Matches 555; Conservative 25; Mismatches 44; Indels 25; Gaps 3;

QY 2 EKLWTVYVGVVWKEATITLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNENYTE 61
 DB 31 EKLWTVYVGVVWKEATITLFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNENYTE 90
 QY 62 FNMWKNMVEQMEDIIISLMDOSLKPCKVLTPLCVTLNCKDV-----NATNTTNDSEG 115
 DB 91 FNMWKNMVEQMEDIIISLMDOSLKPCKVLTPLCVTLNCKDV-----NATNTTNDSEG 150
 QY 116 TMRGEIKNSFNITTSIRDEVOKEVALFYKLDVYXIDNN-NTSYRLISCDTSVITQACP 174
 DB 151 TMRGEIKNSFNITTSIRDEVOKEVALFYKLDVYXIDNN-NTSYRLISCDTSVITQACP 210
 QY 175 KISFEPIPIHYCAPAGFALIKCNDKTFNGKCKPCKNVSTXOCTHGIRPVVSTOLLNGSLA 234
 DB 211 KISFEPIPIHYCAPAGFALIKCNDKTFNGKCKPCKNVSTXOCTHGIRPVVSTOLLNGSLA 270
 QY 235 EEEVVISDNTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIR 276
 DB 271 EEEVVISDNTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIR 330
 QY 277 QAHNINISAKKNDTLKQIVIKAREOFENKTIIVNHSOGDEIYVHSPNCGGEFFYCNST 336
 DB 331 QAHNINISAKKNDTLKQIVIKAREOFENKTIIVNHSOGDEIYVHSPNCGGEFFYCNST 390
 QY 337 QLFNSTWNNNTTSGNNTTEGNITLPCRIRKQIINMVEGKAMVAPPIRGQIRCSNITGL 396
 DB 391 QLFNSTWNNNTTSGNNTTITLPCRIRKQIINMVEGKAMVAPPIRGQIRCSNITGL 450
 QY 397 LITDGGINENGTEIFPPGGGDMRDNMRSELVYKVKVIEPLGVAPTKCRRVVQREKRA 456
 DB 451 LITDGGINENGTEIFPPGGGDMRDNMRSELVYKVKVIEPLGVAPTKCRRVVQREKRA 510
 QY 457 VGIGAVFLGLGAAGSTMGASMTLTVQARLLLSGIVQOQNNLLRAIEAOQMLQLTWG 516

```

Db      511 VGIAGVLELGFAGASTWGASMTLTVAQRLLSGIVQOQNNLRAIEAOQHLLQLTWVG 570
Qy      517 IKQQAVALAVERLIGVQOQLLIGWCSGKILCTTAVPNNASWSKSLDRIANNMTWMEWE 576
Db      571 IKQQAVALAVERLIGVQOQLLIGWCSGKILCTTAVPNNASWSKSLDRIANNMTWMEWE 630
Qy      577 REIDNTSEIYTLIEESONQOEKNEQELLELDKMASLNNWFDITKMLWY 625
Db      631 REINNTSIISLIEESONQOEKNEQELLELDKMASLNNWFDITKMLWY 679

```

RESULT 5

```

ID      09PXW7      PRELIMINARY;      PRT;      846 AA.
AC      09PXW7;
DT      01-MAY-2000 (TReMBLrel. 13, Created)
DT      01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT      01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DT      GP160, envelope glycoprotein.
OC      OS
OX      Human immunodeficiency virus 1.
RN      NCBI_TaxID=11676;
RP      SEQUENCE FROM N.A.
RA      MEDLINE=93189881; PubMed=8446773;
RA      Schmidt-Mayerova H., Gayet O., Guetart N., Bolmont C., Hirsch I.,
RA      Chermann J.C.;
RT      "Characterization of HIV-1-PAR, a macrophage-tropic strain: cell
RT      tropism, virus/cell entry and nucleotide sequence of the envelope
RT      glycoprotein."
RL      Res. Virol. 144:21-26(1993).
DR      InterPro: IPR000328; Env GP41.
DR      InterPro: IPR007777; GP120.
DR      Pfam: PF00516; GP120; 1.
DR      Pfam: PF00517; GP41; 1.
KW      AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ      SEQUENCE 846 AA; 96490 MW; 3D561D5735F90C64 CRC64;

```

```

Query Match      87.7%; Score 2939.5; DB 15; Length 846;
Best Local Similarity 86.2%; Pred. No. 8.1e-232;
Matches 555; Conservative 30; Mismatches 36; Indels 23; Gaps 6;

Qy      2 EKLWTVYVGVPMWKEATTLTFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNTE 61
Db      31 DKLWTVYVGVPMWKEATTLTFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNTE 90
Qy      62 HFNWKNMNVQOMQEDIIISLMDOSLKPVCVLTPLCVTLNCKDV-NATNTNDSGTEMERG 120
Db      91 YFNWKNMNVQOMQEDIIISLMDOSLKPVCVLTPLCVTLNCKDV-NATNTNDSGTEMERG 150
Qy      121 EIKKCSFNITTSIRDEVOKEVALFYKLDVVKIDNNNTSYRLISCDTSYTTQACPKISFEP 180
Db      151 EIKKCSFNITTSIRDEVOKEVALFYKLDVVKIDNNNTSYRLISCDTSYTTQACPKISFEP 210
Qy      181 IPIHYCAPAGPAIKCNDKTFNGKXPCKNVSTXCTGIRPVVSTOLLNGLSLEEEVVI 240
Db      211 IPIHYCAPAGPAIKCNDKTFNGKXPCKNVSTXCTGIRPVVSTOLLNGLSLEEEVVI 270
Qy      241 RSDNFTNNAKTIIVQLKESVEINCTRPNNNG-----AGDIRQAHQNI 282
Db      271 RSDNFTNNAKTIIVQLKESVEIKCTRPNNNTRKSIPIGGRAPFTTGEIIGDIRQAHCTI 330
Qy      283 SRAKMDTLKQIVIKLREQENKTIIVNHSGGDEPEIVMSFNGGGEFFYCNSTQLENST 342
Db      331 SKTMTENFKQIVIKLREQENKTIIVNHSGGDEPEIVMSFNGGGEFFYCNSTQLENST 390
Qy      343 MNNTTBSGNTNTEGTTITLPCRIKOIIMMDEVGKAMVAPPIRGQIRCSNITGLLTRDG 402
Db      391 W-NDTBSGNDTE-RITLTPCRIKOIIMMDEVGKAMVAPPIRGQIRCSNITGLLTRDG 448
Qy      403 G-INENGTEIFRPGGDMDRDMRSELKYKVKVIEPLGVAPTKCKRRVVOREKRAVGIGA 461

```

```

Db      449 GNNNNNGTEIFRPVGGMDRDMRSELKYKVKVIEPLGVAPTKARRRVQREKRAVGIGA 508
Qy      462 VPLGFLGAGASTWGASMTLTVAQRLLSGIVQOQNNLRAIEAOQRMLOLTWVGKQIQ 521
Db      509 MFLGFLGAGASTWGASMTLTVAQRLLSGIVQOQNNLRAIEAOQRLLOLTWVGKQIQ 568
Qy      522 ARVLAVERVYKDOQIILIGWCSGKILCTTAVPNNASWSKSLDRIANNMTWMEWEIDN 581
Db      569 ARVLAVERVYKDOQIILIGWCSGKILCTTAVPNNASWSKSLDRIANNMTWMEWEIDN 627
Qy      582 YTSIYTLIEESONQOEKNEQELLELDKMASLNNWFDITKMLWY 625
Db      628 YTNLIYTLIEESONQOEKNEQELLELDKMASLNNWFDITKMLWY 671

```

RESULT 6

```

ID      003805      PRELIMINARY;      PRT;      855 AA.
AC      003805;
DT      01-NOV-1996 (TReMBLrel. 01, Created)
DT      01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT      01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE      Env protein.
GN      ENV.
OS      Human immunodeficiency virus 1.
OC      Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX      NCBI_TaxID=11676;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Reitz M., Popovic M., Gartner S., Gallo R.C., Reed-Connole E.,
RA      Beaver B.;
RL      Submitted (JUN-1991) to the EMBL/GenBank/DBJ databases.
DR      EMBL: M68893; AAA44191.1; -.
DR      InterPro: IPR000328; Env GP41.
DR      InterPro: IPR007777; GP120.
DR      Pfam: PF00516; GP120; 1.
DR      Pfam: PF00517; GP41; 1.
SQ      SEQUENCE 855 AA; 96988 MW; EC631A3B84180C8D CRC64;

```

```

Query Match      87.7%; Score 2938.5; DB 15; Length 855;
Best Local Similarity 85.1%; Pred. No. 9.9e-232;
Matches 552; Conservative 27; Mismatches 45; Indels 25; Gaps 3;

Qy      2 EKLWTVYVGVPMWKEATTLTFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNTE 61
Db      31 DKLWTVYVGVPMWKEATTLTFCASDAKAYDTEVHNWATHACVPTDPNPOEVLNTE 90
Qy      62 HFNWKNMNVQOMQEDIIISLMDOSLKPVCVLTPLCVTLNCKDV-----NATNTNDSG 115
Db      91 NFNWKNMNVQOMQEDIIISLMDOSLKPVCVLTPLCVTLNCKDV-----NATNTNDSG 150
Qy      116 TMBERGEIKNCSFNITTSIRDEVOKEVALFYKLDVVKIDNN-NTSYRLISCDTSYTTQACP 174
Db      151 MVGGEIKNCSFNITTSIRDEVOKEVALFYKLDIAPIDNNNSNRYSRLISNTSVITQACP 210
Qy      175 KISPEPIHYCAPAGPAIKCNDKTFNGKXPCKNVSTXCTGIRPVVSTOLLNGLSLE 234
Db      211 KVSPEPIHYCAPAGPAIKCNDKTFNGKXPCKNVSTXCTGIRPVVSTOLLNGLSLE 270
Qy      235 EEEVIRSDNFTNNAKTIIVQLKESVEINCTRPNNNG-----AGDIR 276
Db      271 EEEVIRSDNFTNNAKTIIVQLKESVEINCTRPNNNTRKSIPIGGRAPFTTGEIIGDIR 330
Qy      277 QAHNISRKAMDTLKQIVIKLREQENKTIIVNHSGGDEPEIVMSFNGGGEFFYCNST 336
Db      331 QAHNISRKAMDTLKQIVIKLREQENKTIIVNHSGGDEPEIVMSFNGGGEFFYCNST 390
Qy      337 QLFNSTNNNTBSGNTNTEGTTITLPCRIKOIIMMDEVGKAMVAPPIRGQIRCSNITGL 396
Db      391 QLFNSTNNNTBSGNTNTEGTTITLPCRIKOIIMMDEVGKAMVAPPIRGQIRCSNITGL 450
Qy      397 LLTRDGINENGTEIFRPGGDMDRDMRSELKYKVKVIEPLGVAPTKCKRRVVOREKRA 456

```

```
Db 451 LITRGGPEPDNTEVFRPGGDMRDWRSELYKYKVKLEPGVAPTKAKRRVVOREKA 510
Qy 457 VGIQAVFLGFLGAAGSTMGASMTLTVQARLLLSGIVQOONNLLAIEAOQMLDLTWG 516
Db 511 VGIQAVFLGFLGAAGSTMGASMTLTVQARLLLSGIVQOONNLLAIEAOQMLDLTWG 570
Qy 517 IKQLQARVLAVERYLADQDLGIGWCGSKLICCTAVPNNASVSNKSLDRINNNMTWME 576
Db 571 IKQLQARVLAVERYLADQDLGIGWCGSKLICCTAVPNNASVSNKSLDRINNNMTWME 630
Qy 577 REIDNVTSEIYTLIESQNOQEKNOBELLELDKMASLNNWPDITKWLWY 625
Db 631 REINNYTSIYSLIESQNOQEKNOBELLELDKMASLNNWPDITKWLWY 679

RESULT 7
ID 040222 PRELIMINARY; PRT; 854 AA.
040222;
01-JAN-1998 (TREMBlrel. 05, Created)
01-JAN-1998 (TREMBlrel. 05, Last sequence update)
01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Env polyprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AD8;
RX MEDLINE=96432129; PubMed=8835195;
RA Theodore T.S., Englund G., Buckler-White A., Buckler C.E.,
RA Martin M.A., Peden K.W.;
RA "Construction and characterization of a stable full-length macrophage-
RT tropic HIV type 1 molecular clone that directs the production of high
RT titers of progeny virions."
RL AIDS Res. Hum. Retroviruses 12:191-194 (1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AD8;
RA Theodore T.S., Englund G., Buckler-White A., Buckler C.E.,
RA Martin M.A., Peden K.W.C.;
RA Submitted (MAY-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF004394; AAB64170.1; -
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
SQ AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 854 AA; 97291 MW; 06C45869103C6C12 CRC64;

Query Match 87.6%; Score 2936.5; DB 15; Length 854;
Best Local Similarity 86.2%; Pred. No. 1.4e-231;
Matches 562; Conservative 18; Mismatches 43; Indels 29; Gaps 8;

Qy 1 VEKLVATVYVYGVVWKEATTTTFCASDAKAYDTEVHNVATTAACVPTDPNPOEVLENT 60
Db 30 VENLVATVYVYGVVWKEATTTTFCASDAKAYDTEVHNVATTAACVPTDPNPOEVLENT 89
Qy 61 EHFNNKNNMVEQMOEDIIISLDOSLKPCKVLTPLCVTLNCKDV-NATVTTDSSEGTMR 119
Db 90 EHFNNKNNMVEQMOEDIIISLDOSLKPCKVLTPLCVTLNCKDV-NATVTTDSSEGTMR 147
Qy 120 GEIKKCSFNITTSIRDEVQKEYALFYKLDVYXIDNNNTSYRLISCDSVTITQACPKISFE 179
Db 148 GEIKKCSFNITTSIRDEVQKEYALFYKLDVYXIDNNNTSYRLISCDSVTITQACPKISFE 207
Qy 180 PPIHYCAPGFAILKCNDRKNGKXPCKNVSTXOCTHGIRPVVSTOLLNGLAEEVYR 239
Db 208 PPIHYCAPGFAILKCNDRKNGKXPCKNVSTXOCTHGIRPVVSTOLLNGLAEEVYR 267
Qy 240 IRSDFNTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRQAHCN 281
```

```
Db 268 IRSSNFTDNNAKTIIVOLKESVEINCTRPNNNTRKSIHIGPGRAFYTTGDIIGDIRQAHCN 327
Qy 282 ISRAKNNNTLQOIYKLEQOF-ENKTIYFNHSSGGDEIIVHSPFCGSEFPFNCSTOLF 340
Db 328 ISRTKNNNTLQOIYKLEQOF-ENKTIYFNHSSGGDEIIVHSPFCGSEFPFNCSTOLF 387
Qy 341 STWNN-----NTEGSNTEGN-TITLPCRIRKQIINMMQEVGKAMPAPPIRGQIRCSNITG 395
Db 368 STWNNNGTWNLQSGNTEGNITLPCRIRKQIINMMQEVGKAMPAPPIRGQIRCSNITG 447
Qy 366 LLITRDGGINENG-TEIFRPGGDMRDWRSELYKYKVKLEPGVAPTKAKRRVVOREK 454
Db 448 LLITRDGGINENNDETEFRPGGDMRDWRSELYKYKVKLEPGVAPTKAKRRVVOREK 507
Qy 455 RAVG-IGAVFLGFLGAAGSTMGASMTLTVQARLLLSGIVQOONNLLAIEAOQMLDLTW 513
Db 508 RAVGTIGAVFLGFLGAAGSTMGASMTLTVQARLLLSGIVQOONNLLAIEAOQMLDLTW 567
Qy 514 VGIQQLQARVLAVERYLADQDLGIGWCGSKLICCTAVPNNASVSNKSLDRINNNMTW 573
Db 568 VGIQQLQARVLAVERYLADQDLGIGWCGSKLICCTAVPNNASVSNKSLDRINNNMTW 627
Qy 574 EMEREIDNVTSEIYTLIESQNOQEKNOBELLELDKMASLNNWPDITKWLWY 625
Db 628 EMEREIDNVTSEIYTLIESQNOQEKNOBELLELDKMASLNNWPDITKWLWY 679

RESULT 8
ID 003806 PRELIMINARY; PRT; 838 AA.
003806;
01-NOV-1996 (TREMBlrel. 01, Created)
01-NOV-1996 (TREMBlrel. 01, Last sequence update)
01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Env protein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Reitz M., Popovic M., Gartner S., Gallo R.C., Reed-Connole E.,
RA Beaver B.;
RA Submitted (JUN-1991) to the EMBL/Genbank/DBJ databases.
DR EMBL; M68894; AAA44196.1; -
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
SQ SEQUENCE 838 AA; 95329 MW; 11044B1CA10CFB4D CRC64;

Query Match 87.6%; Score 2935.5; DB 15; Length 838;
Best Local Similarity 85.8%; Pred. No. 1.7e-231;
Matches 551; Conservative 24; Mismatches 40; Indels 27; Gaps 2;

Qy 2 EKLWATVYVYGVVWKEATTTTFCASDAKAYDTEVHNVATTAACVPTDPNPOEVLENT 61
Db 31 EKLWATVYVYGVVWKEATTTTFCASDAKAYDTEVHNVATTAACVPTDPNPOEVLENT 90
Qy 62 HFNMMKNNMVEQMOEDIIISLDOSLKPCKVLTPLCVTLNCKDV-NATVTTDSSEGTMR 121
Db 91 HFNMMKNNMVEQMOEDIIISLDOSLKPCKVLTPLCVTLNCKDV-NATVTTDSSEGTMR 141
Qy 122 IKKCSFNITTSIRDEVQKEYALFYKLDVYXIDNNNTSYRLISCDSVTITQACPKISFE 181
Db 142 MKNCSFNITTSIRDEVQKEYALFYKLDVYXIDNNNTSYRLISCDSVTITQACPKISFE 201
Qy 182 PPIHYCAPGFAILKCNDRKNGKXPCKNVSTXOCTHGIRPVVSTOLLNGLAEEVYR 241
Db 202 PPIHYCAPGFAILKCNDRKNGKXPCKNVSTXOCTHGIRPVVSTOLLNGLAEEVYR 261
Qy 242 SDFNTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRQAHCN 283
```



```

Db 262 SENFTNNAKTIIVQINSEVEINCTRPNNTRKSIINGRAFYTTGELIGDIRQAHCNLS 321
Qy 284 RAKNDLTKQIVIKLREGFENKTIIVFNHSSGGDEIIVNHSFNGGFEFFCNSTOLFENSTW 343
Db 322 RAKNDLTKQIVIKLREGFENKTIIVFNHSSGGDEIIVNHSFNGGFEFFCNSTOLFENSTW 381
Qy 344 NNNTGSSNTEGNTITLPCRIRKQIINNMOEKGAMVAPPIRGQIRCSNITGLLTDGG 403
Db 382 NVTESNNTVENNTITLPCRIRKQIINNMOEKGAMVAPPIRGQIRCSNITGLLTDGG 441
Qy 404 INENGTETFRPGGDMNRSELYKVKYKIEPLGVAPTKCRVVOREKRAVGIG 463
Db 442 PEDKTEVFREPGGDMNRSELYKVKYKIEPLGVAPTKCRVVOREKRAVGIG 501
Qy 464 LGFLGAGSTMGASMTLTVQARLLSGIVQOONNLRAIEAOQRLQITVWGIKOL 523
Db 502 LGFLGAGSTMGASMTLTVQARLLSGIVQOONNLRAIEAOQRLQITVWGIKOL 561
Qy 524 VLAVERYLGDQQLIGWCSGKLICTTAVPNNASWSNKSIDRIINNMTWMEEREID 583
Db 562 VLAVERYLGDQQLIGWCSGKLICTTAVPNNASWSNKSIDRIINNMTWMEEREID 621
Qy 584 SEIYTLIEESQNOQEKNEQELBELDKWASLMMNFDITKWLWY 625
Db 622 SIYSLIEESQNOQEKNEQELBELDKWASLMMNFDITKWLWY 663

```

RESULT 9

```

Qy 077368 PRELIMINARY; PRT; 849 AA.
ID 077368
AC 077368;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMblrel. 22, Last annotation update)
DE Env polypeptide.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Buckler-White A., Theodore T.
RL Submitted (OCT-1992) to the EMBL/Genbank/DBJ databases.
DR EMBL; M38430; AAA44073.1; -.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR AIDS; Coat protein; Glycoprotein; Polypeptide; Transmembrane.
SQ SEQUENCE 849 AA; 97166 MW; 8DBEGBA86AA92467 CRC64;

```

```

Query Match 87.5%; Score 2931; DB 15; Length 849;
Best Local Similarity 86.4%; Pred. No. 4e-231;
Matches 557; Conservative 28; Mismatches 38; Indels 22; Gaps 5;
Qy 2 EKLWTVYGVVWPKKATTLTFCASDAKAYDTEVHNWATHACVPTDPNQEVLLENTE 61
Db 31 EKLWTVYGVVWPKKATTLTFCASDAKAYDTEVHNWATHACVPTDPNQEVLLENTE 90
Qy 62 HNNMMKNMVEQOMEDIISLMDQSLKPCVKLTPLCVTLNCKD--VNATNTNDSGTEMERG 120
Db 91 NFNMMKNMVEQOMEDIISLMDQSLKPCVKLTPLCVTLNCKD--VNATNTNDSGTEMERG 150
Qy 121 EIKNCSFNITTSIRDEVOKEYALFYKLDVXKIDNNNTSYRLISCDTSYTTQACPKISFE 180
Db 151 EIKNCSFNITTSIRDEVOKEYALFYKLDVXKIDNNNTSYRLISCDTSYTTQACPKISFE 210
Qy 181 IPIHYCAPAGFALLKCNCKTNGKKPKCNVSTXOCTHGRIPVSTOLLNLSLSEEVY 240
Db 211 IPIHYCAPAGFALLKCNCKTNGKKPKCNVSTXOCTHGRIPVSTOLLNLSLSEEVY 270
Qy 241 RSDNFTNNAKTIIVQINSEVEINCTRPN-----NNGAG-----DIRQAHCN 282

```

```

Db 271 RSDNFTNNAKTIIVQINSEVEINCTRPNIYRKGRIRIHGPGRAPHTTROIENIROAHCN 330
Qy 283 SRKNDLTKQIVIKLREGFENKTIIVFNHSSGGDEIIVNHSFNGGFEFFCNSTOLFENST 342
Db 331 SRKNDLTKQIVIKLREGFENKTIIVFNHSSGGDEIIVNHSFNGGFEFFCNSTOLFENST 390
Qy 343 MN--NTGSSNTEGNTITLPCRIRKQIINNMOEKGAMVAPPIRGQIRCSNITGLLTD 400
Db 391 MNSTGKSSNNTNEN--ITLPCRIRKQIINNMOEKGAMVAPPIRGQIRCSNITGLLTD 449
Qy 401 DGINENGTETFRPGGDMNRSELYKVKYKIEPLGVAPTKCRVVOREKRAVGIG 460
Db 450 DGINENGTETFRPGGDMNRSELYKVKYKIEPLGVAPTKCRVVOREKRAVGIG 509
Qy 461 AVFLGFLGAGSTMGASMTLTVQARLLSGIVQOONNLRAIEAOQRLQITVWGIKOL 520
Db 510 AVFLGFLGAGSTMGASMTLTVQARLLSGIVQOONNLRAIEAOQRLQITVWGIKOL 569
Qy 521 QARVLAVERYLGDQQLIGWCSGKLICTTAVPNNASWSNKSIDRIINNMTWMEEREID 580
Db 570 QARVLAVERYLGDQQLIGWCSGKLICTTAVPNNASWSNKSIDRIINNMTWMEEREID 629
Qy 581 NYTSEIYTLIEESQNOQEKNEQELBELDKWASLMMNFDITKWLWY 625
Db 630 NYTSEIYTLIEESQNOQEKNEQELBELDKWASLMMNFDITKWLWY 674

```

RESULT 10

```

Qy 092762 PRELIMINARY; PRT; 854 AA.
ID 092762
AC 092762;
DT 01-NOV-1998 (TREMblrel. 08, Created)
DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
DT 01-OCT-2002 (TREMblrel. 22, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=SFMS6.3;
RC MEDLINE=98178716; PubMed=9519894;
RA McCutcheon F.E., Sanders-Buell E., Salminen M.O., Carr J.K.,
  Sheppard W.H.;
RT "Diversity of the human immunodeficiency virus type 1 envelope
  glycoprotein in San Francisco Men's Health Study participants.";
RL AIDS Res. Hum. Retroviruses 14:329-337(1998).
DR EMBL; AF025754; AAC40591.1; -.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR AIDS; Coat protein; Glycoprotein; Polypeptide; Transmembrane.
SQ SEQUENCE 854 AA; 97287 MW; AFD8A82B62C5B27D CRC64;

```

```

Query Match 87.4%; Score 2929; DB 15; Length 854;
Best Local Similarity 85.4%; Pred. No. 5.9e-231;
Matches 554; Conservative 30; Mismatches 39; Indels 26; Gaps 6;
Qy 2 EKLWTVYGVVWPKKATTLTFCASDAKAYDTEVHNWATHACVPTDPNQEVLLENTE 61
Db 32 EKLWTVYGVVWPKKATTLTFCASDAKAYDTEVHNWATHACVPTDPNQEVLLENTE 91
Qy 62 HNNMMKNMVEQOMEDIISLMDQSLKPCVKLTPLCVTLNCKD--VNATNTNDSGTEMER 119
Db 92 NFNMMKNMVEQOMEDIISLMDQSLKPCVKLTPLCVTLNCKD--VNATNTNDSGTEMER 151
Qy 120 GEIKNCSFNITTSIRDEVOKEYALFYKLDVXKIDNNNTSYRLISCDTSYTTQACPKISFE 179
Db 152 GEIKNCSFNITTSIRDEVOKEYALFYKLDVXKIDNNNTSYRLISCDTSYTTQACPKISFE 211

```


QY 180 PIPHYCAPAGFALLKCNKDTFNCKPCPCNNSTXOCTHGIRPVVSTOLLNGLAEEVY 239
 Db 212 PIPHYCAPAGFALLKCNKDTFNCKPCPCNNSTXOCTHGIRPVVSTOLLNGLAEEVY 271
 QY 240 IRSDFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRQACHN 281
 Db 272 IRSDFTNNAKTIIVOLKESVEINCTRPNNNRKSIPIGPGAFYTTGELIIGNIRQACHN 331
 QY 282 ISRAKNDTLKOIVIKLREOFENKTIIVFNHSGGDEPEIYHMSFNGCGEPFYCNSTQLFNS 341
 Db 332 ISRAKNDTLKOIVIKLREOFENKTIIVFNHSGGDEPEIYHMSFNGCGEPFYCNSTQLFNS 391
 QY 342 TAWN-NTEGSNTEGN-TITLPCRIKOIINMOWEYKAMAPPIRGQIRCSNITGLLLT 399
 Db 392 TAWNTEGSNTEGN-TITLPCRIKOIINMOWEYKAMAPPIRGQIRCSNITGLLLT 451
 QY 400 RDGGINENGT---EIFRPGGDMRNRSELYKYKVVNTEIPELVAPTCKRVRVQREKRA 456
 Db 452 RDGGINENGT---EIFRPGGDMRNRSELYKYKVVNTEIPELVAPTCKRVRVQREKRA 510
 QY 457 VGIGAVFPGFLGAAGSTWGAASMTLTVQARLLSGIVQOQNNLLRAIEAQOBLQITWVG 516
 Db 511 AGIGAVFPGFLGAAGSTWGAASMTLTVQARLLSGIVQOQNNLLRAIEAQOBLQITWVG 570
 QY 517 IKOLQARVLAVERLYGDOQLLGIGWCGSGKLICTTAVPWNASWSKSLDRINNNMTWME 576
 Db 571 IKOLQARVLAVERLYGDOQLLGIGWCGSGKLICTTAVPWNASWSKSLDRINNNMTWME 630
 QY 577 REINDYTSIYTLIESONQOEKNEBELDKKASLMMNPFITKMLY 625
 Db 631 REINDYTSIYTLIESONQOEKNEBELDKKASLMMNPFITKMLY 679

RESULT 11

Q77989 PRELIMINARY; PRT; 863 AA.
 AC Q77989;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
 DE Envelope glycoprotein.
 GN ENV.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SP128A;
 MEDLINE=91056585; PubMed=2243391;
 RT Liu Z.-Q., Mood C., Levy J.A., Cheng-Mayer C.;
 RT "The viral envelope gene is involved in macrophage tropism of a human
 RT immunodeficiency virus type 1 strain isolated from brain tissue.";
 RL J. Virol. 64:6148-6153 (1990).
 DR EMBL; M95292; AAA44331.1; -;
 DR InterPro; IPR000328; Env_GP41.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00517; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 DR AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
 KW SEQUENCE 863 AA; 98222 MW; 0157397FAB84C5D7 CRC64;
 SQ

Query Match 87.4%; Score 2928; DB 15; Length 863;
 Best Local Similarity 84.4%; Pred. No. 7, 2e-231;
 Matches 557; Conservative 25; Mismatches 40; Indels 38; Gaps 7;

QY 2 EKLWTVYGVVPMWKEATTTLLFCASDAKAYDTEVHNWATACVPTDPNPQEVYLENVT 61
 Db 31 EKLWTVYGVVPMWKEATTTLLFCASDAKAYDTEVHNWATACVPTDPNPQEVYLENVT 90
 QY 62 HNNMKNNMVEOMODIISLNDOSLKPCVKLTPLCVTLNCD---VNAITNDSGEM 117
 Db 91 HNNMKNNMVEOMODIISLNDOSLKPCVKLTPLCVTLNCD---VNAITNDSGEM 149

QY 118 -----ERGEINQCSFNITTSIRDEVQEVAFYKLDVXIDNNNTSYRLICDTSVITQA 172
 Db 150 SLSGETGEIKQCSFNITTSIRDKQKQEVAFYKLDVPIVNNNTYRLINCNSTITQA 209
 QY 173 CPKISFEPIPIHYCAPAGFALLKCNKDTFNCKPCPCNNSTXOCTHGIRPVVSTOLLNGLS 232
 Db 210 CPKISFEPIPIHYCAPAGFALLKCNKDTFNCKPCPCNNSTXOCTHGIRPVVSTOLLNGLS 269
 QY 223 LAEEVVRSDNFTNNAKTIIVOLKESVEINCTRPNNN-----GA--GD 274
 Db 270 LAEEVVRSDNFTNNAKTIIVOLKESVEINCTRPNNNRKSIPIGPGAFYTTGELIIGNIRQACHN 329
 QY 275 IROACHNISRAKNDTLKOIVIKLREOFENKTIIVFNHSGGDEPEIYHMSFNGCGEPFYCN 334
 Db 330 IROACHNISRAKNDTLKOIVIKLREOFENKTIIVFNHSGGDEPEIYHMSFNGCGEPFYCN 389
 QY 335 STOLFNSNTEGN-TITLPCRIKOIINMOWEYKAMAPPIRGQIRCSNITGLLLT 385
 Db 390 STOLFNSNTEGN-TITLPCRIKOIINMOWEYKAMAPPIRGQIRCSNITGLLLT 448
 QY 386 QIRCSNITGLLLTRDGINENGT---EIFRPGGDMRNRSELYKYKVVNTEIPELVAPTCKRVRVQREKRA 445
 Db 449 QIRCSNITGLLLTRDGINENGT---EIFRPGGDMRNRSELYKYKVVNTEIPELVAPTCKRVRVQREKRA 508
 QY 446 KRRVQREKRAVIGAVFPGFLGAAGSTWGAASMTLTVQARLLSGIVQOQNNLLRAIEA 505
 Db 509 KRRVQREKRAVIGAVFPGFLGAAGSTWGAASMTLTVQARLLSGIVQOQNNLLRAIEA 568
 QY 506 QORMLQITWVGIKOLQARVLAVERLYGDOQLLGIGWCGSGKLICTTAVPWNASWSKSLDR 565
 Db 569 QORMLQITWVGIKOLQARVLAVERLYGDOQLLGIGWCGSGKLICTTAVPWNASWSKSLDR 628
 QY 566 IMNNMTWMEERREINDYTSIYTLIESONQOEKNEBELDKKASLMMNPFITKMLY 625
 Db 629 IMNNMTWMEERREINDYTSIYTLIESONQOEKNEBELDKKASLMMNPFITKMLY 688

RESULT 12

Q003811 PRELIMINARY; PRT; 853 AA.
 AC Q003811;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Surface envelope glycoprotein.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ADA;
 MEDLINE=91195299; PubMed=2014229;
 RT Westervelt P., Gendelman H.E., Ratner L.;
 RT "Identification of a determinant within the human immunodeficiency
 RT virus 1 surface envelope glycoprotein critical for productive
 RT infection of primary monocytes.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:3097-3101 (1991).
 DR EMBL; M60472; AAA45065.1; -;
 DR InterPro; IPR000328; Env_GP41.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 DR Envelope protein.
 KW SEQUENCE 853 AA; 96998 MW; AD2AP21E2B06AD78 CRC64;
 SQ

Query Match 87.3%; Score 2926; DB 15; Length 853;
 Best Local Similarity 85.4%; Pred. No. 1e-230;
 Matches 556; Conservative 25; Mismatches 42; Indels 28; Gaps 7;

QY 1 VEKLMWTVYGVVPMWKEATTTLLFCASDAKAYDTEVHNWATACVPTDPNPQEVYLENVT 60
 Db 30 VEKLMWTVYGVVPMWKEATTTLLFCASDAKAYDTEVHNWATACVPTDPNPQEVYLENVT 89

QY 61 EHFMMKNNVQEQEDIIISLMDQSLKPCVLTPLCTLANCKDY-NATNTNDSEGTMER 119
 Db 90 ENFMNMKNWQEQHEDIIISLMDQSLKPCVLTPLCTLANCKDY-NATNTNDSEGTMER 147
 QY 120 GEIKNSFNITTSIRDEYQKEYALFYKLDVYXIDNNNTSYRLISCDTSVITQACPISFE 179
 Db 148 GEIKNSFNITTSIRDKVKDYALFYRLDVPIDNDNTSYRLINCNSTITQACPISFE 207
 QY 180 PIPIHYCAPGAFALIKCKDKTFNGKXPCKNVSTQCHGRIPVYSTOLLNGSLAESEV 239
 Db 208 PIPIHYCAPGAFALIKCKDKTFNGKXPCKNVSTQCHGRIPVYSTOLLNGSLAESEV 267
 QY 240 IRSNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRAOHCN 281
 Db 268 IRSNFTNNAKTIIVOLKESVEINCTRPNNNTKRSIHGPRAYTTGEIIGDIRAOHCN 327
 QY 282 ISRAKNDTLKQIYIKLREOF-ENKTIYFNHSSGDEPEIYVMSFNGGGEFFYCNSTOLFEN 340
 Db 328 ISRTKNNNTLQIATKKEQFGNNKTIYFNQSGGDEPEIYVMSFNGGGEFFYCNSTOLFEN 387
 QY 341 STANN-----NTEGSNTEGN-TITLPCRIKOIINMOEYKAMVAPPIRGQIRCSNITG 395
 Db 388 STANNFTNNTLQSNTEGNITLPCRIKOIINMOEYKAMVAPPIRGQIRCSNITG 447
 QY 396 LILTRDGINENGTETFRPGGDMRDNRSELKYKVKIEPLGVAAPTCKGRVOREKR 455
 Db 448 LILTRDGINENGTETFRPGGDMRDNRSELKYKVKIEPLGVAAPTCKGRVOREKR 507
 QY 456 AVG-IGAVFLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOQNNLLRAIEAQRMLOLTV 514
 Db 508 AVGTGAMFLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOQNNLLRAIEAQRMLOLTV 567
 QY 515 WGIQOLQARVLAVERLYLDQOLLGIWCGSGKLICTTAVPMNASNSKSLDRIINNTWME 574
 Db 568 WGIQOLQARVLAVERLYLDQOLLGIWCGSGKLICTTAVPMNASNSKSLDRIINNTWME 627
 QY 575 WEREIDVYSEIYTLIESQNOEQKNEQELLEDKMASLNNPDTIKMLY 625
 Db 628 WEREIDVYSEIYTLIESQNOEQKNEQELLEDKMASLNNPDTIKMLY 678

RESULT 13

OS 056110 PRELIMINARY; PRT; 851 AA.
 AC 056110;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 Db Envelope glycoprotein.

ENV.

OS Human immunodeficiency virus 1.
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.
 NC NCB1_TaxID=11676;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=SPMHS8.1;
 RX MEDLINE=98178716; PubMed=9519894;
 RA McCutchan F.E., Sanders-Buell E., Salminen M.O., Carr J.K.,
 RA Sheppard W.H.;

RT "Diversity of the human immunodeficiency virus type 1 envelope
 RT glycoprotein in San Francisco Men's Health Study participants.",
 RL AIDS Res. Hum. Retroviruses 14:329-337(1998).
 DR EMBL, AF025756; AAC40593.1;
 DR InterPro; IPR000328; Env_GP41.
 DR InterPro; IPR000173; GAP_dhdrogenase.
 DR Pfam; PF00516; GP120.1.
 DR Pfam; PF00517; GP41.1.
 DR PROSITE; PS00071; GAPDH.1.
 KW AIDS; Coat protein; Glycoprotein; Polypeptide; Transmembrane.
 SQ SEQUENCE 851 AA; 96516 MW; 1CDAC83EB6464531 CRC64;

Query Match 87.3%; Score 2924; DB 15; Length 851;

Best Local Similarity 85.6%; Pred. No. 1.5e-230;
 Matches 555; Conservative 27; Mismatches 40; Indels 26; Gaps 5;

QY 2 EKLMTVYVGVVWPKKATITLFCASDAKAYDEVHVMMAHACVPPDPDQEVLENYTE 61
 Db 31 DKLMVTVYVGVVWPKKATITLFCASDAKAYDEVHVMMAHACVPPDPDQEVLENYTE 90
 QY 62 HFMNMKNWQEQEDIIISLMDQSLKPCVLTPLCTLANCKDY-NATNTNDSEGTMER 120
 Db 91 HFMNMKNWQEQHEDIIISLMDQSLKPCVLTPLCTLANCKDY-NATNTNDSEGTMER 150
 QY 121 EIKNSFNITTSIRDEYQKEYALFYKLDVYXIDNNNTSYRLISCDTSVITQACPISFE 180
 Db 151 EIKNSFNITTSIRDKVKDYALFYRLDVPIDNDNTSYRLINCNSTITQACPISFE 209
 QY 181 PIPIHYCAPGAFALIKCKDKTFNGKXPCKNVSTQCHGRIPVYSTOLLNGSLAESEV 240
 Db 210 PIPIHYCAPGAFALIKCKDKTFNGKXPCKNVSTQCHGRIPVYSTOLLNGSLAESEV 269
 QY 241 IRSNFTNNAKTIIVOLKESVEINCTRPNNNG-----AGDIRAOHCN 282
 Db 270 IRSNFTNNAKTIIVOLKESVEINCTRPNNNTKRSIHGPRAYTTGEIIGDIRAOHCN 329
 QY 283 ISRAKNDTLKQIYIKLREOF-ENKTIYFNHSSGDEPEIYVMSFNGGGEFFYCNSTOLFEN 342
 Db 330 ISRAKNDTLKQIYIKLREOF-ENKTIYFNHSSGDEPEIYVMSFNGGGEFFYCNSTOLFEN 389
 QY 343 W-----NTEGSNTEGN-TITLPCRIKOIINMOEYKAMVAPPIRGQIRCSNITGL 397
 Db 390 WSVSTNNDTEGNNT-GENITLPCRIKOIINMOEYKAMVAPPIRGQIRCSNITGL 448
 QY 398 LTRDGINENGTETFRPGGDMRDNRSELKYKVKIEPLGVAAPTCKGRVOREKR 457
 Db 449 LTRDGINENGTETFRPGGDMRDNRSELKYKVKIEPLGVAAPTCKGRVOREKR 508
 QY 458 GIGAVFLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOQNNLLRAIEAQRMLOLTVGI 517
 Db 509 GIGAVFLGFLGAAGSTWGAASMTLTVQARLLLSGIVQOQNNLLRAIEAQRMLOLTVGI 568
 QY 518 KOLQARVLAVERLYLDQOLLGIWCGSGKLICTTAVPMNASNSKSLDRIINNTWME 577
 Db 569 KOLQARVLAVERLYLDQOLLGIWCGSGKLICTTAVPMNASNSKSLDRIINNTWME 628
 QY 578 EIDVYSEIYTLIESQNOEQKNEQELLEDKMASLNNPDTIKMLY 625
 Db 629 EIDVYSEIYTLIESQNOEQKNEQELLEDKMASLNNPDTIKMLY 676

RESULT 14

OS 08AR21 PRELIMINARY; PRT; 847 AA.
 AC 08AR21;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 Db Envelope glycoprotein.

OS Human immunodeficiency virus 1.
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.
 NC NCB1_TaxID=11676;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=SHIVSF162P3.2;
 RX Gao F.;

RL

RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF536757; AA005642.1;
 SQ SEQUENCE 847 AA; 95879 MW; 9CF863B97B3BB54C CRC64;

Query Match 87.0%; Score 2917; DB 15; Length 847;
 Best Local Similarity 85.9%; Pred. No. 5.6e-230;
 Matches 554; Conservative 27; Mismatches 42; Indels 22; Gaps 5;
 QY 1 VEKLMVTVYVGVVWPKKATITLFCASDAKAYDEVHVMMAHACVPPDPDQEVLENYTE 60

THIS PAGE BLANK (US:TO)